DOCUMENT RESUME

ED 264 014 PS 015 487

AUTHOR Mann, Barbara Kay; Sherman, Wendy S.
TITLE A Study of the Predictive Validity of the Dallas Preschool Screening Test; and The Predictive Validity of the Dallas Preschool Screening Test using the Hillsborough Early Rating Scale as Criterion.
INSTITUTION Florida Educational Research and Development Council, Inc., Sanibel.
PUB DATE 84
NOTE 80p.
AVAILABLE FROM FERDC, P.O. Box 506, Sanibel, FL 33957 ($3.00; 10 discount on orders of 5 or more copies. Orders of less than $20.00 must be accompanied by a check or money order.)
PUB TYPE Reports - Research/Technical (143)
JOURNAL CIT Florida Educational Research and Development Council, Inc. Research Bulletin; v18 n1 Sum 1984

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Early Childhood Education; *Kindergarten Children; Learning Problems; *Predictive Validity; *Preschool Children; Preschool Tests; Rating Scales; *School Readiness; *Screening Tests
IDENTIFIERS Comprehensive Tests of Basic Skills; *Dallas Preschool Screening Test; Developmentally Delayed; Hillsborough Early Rating Scale

ABSTRACT

The purpose of these studies was to provide information on the Dallas Preschool Screening Test (DPST). While the DPST is widely used in Texas, little research on the instrument exists. Study 1 provides data on the predictive validity of the DPST for total and component scores of the Comprehensive Test of Basic Skills (CTBS). It was hypothesized that the DPST total and partial scores would significantly correlate with the CTBS Reading Total as would a subset of the DPST subtest scores with the CTBS Reading Total score. Subjects were 107 children with Fall, 1981 DPST scores from the Center for Children with Special Learning Needs plus Spring, 1983 CTBS scores from Hillsborough County Public Schools. Although results indicated some significant correlations, the size of the correlations suggested that they have little predictive value. It is recommended that the DPST be used in conjunction with some other measure. A second study of the predictive validity of the DPST investigated the relationship between preschool performance, as measured by the DPST, and kindergarten success, as measured by the Hillsborough Early Rating Scale (HERS); the relationship betwen the CTBS and the HERS also was explored. Subjects were 101 kindergarten children. Results indicated that the DPST had a limited predictive relationship with the HERS and that the CTBS had only a minimal predictive relationship with the HERS. Results suggest that the use of the DPST as a predictor of later school success is questionable. (RH)

FLORIDA EDUCATIONAL
RESEARCH AND DEVELOPMENT COUNCIL, INC.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

V This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Fla. Educ. Res. & Dev Council, Inc.
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED264014

RESEARCH BULLETIN

A STUDY OF THE PREDICTIVE VALIDITY OF THE DALLAS PRESCHOOL SCREENING TEST

by
Barbara Kay Mann

THE PREDICTIVE VALIDTY OF THE DALLAS PRESCHOOL SCREENING TEST USING THE HILLSBOROUGH EARLY RATING SCALE AS CRITERION

by
Wendy S. Sherman

PS 015 487

Volume 18 Summer, 1984 Number 1

FLORIDA EDUCATIONAL
RESEARCH AND DEVELOPMENT COUNCIL

RESEARCH BULLENTIN

A STUDY OF THE PREDICTIVE VALIDITY OF THE DALLAS PRESCHOOL SCREENING TEST

by
Barbara Kay Mann

THE PREDICTIVE VALIDITY OF THE DALLAS PRESCHOOL SCREENING TEST USING THE HILLSBOROUGH EARLY RATING SCALE AS CRITERION

by
Wendy S. Sherman

PUBLISHED BY
Florida Educational Research
and
Development Council, Inc.

Additional copies of this book may be obtained from:
FERDC
P.O. Box 506
Sanibel, Florida 33957

Individual Copies $3.00
Annual Subscription (4 Issues) $10.00
10% Discount on an order of 5 or more
Order of less than $20.00 must be accompanied by a check or money order.

F.E.R.D.C. NOTES ON THIS EDITION

The constant search for predictive instruments to assist educators to make decisions about very young children prompted Florida Educational Research and Development Council, Inc. to fund the research and publish the results of two studies. These two studies are intended to assist those responsible for testing and measurement by providing some additional data about the reliability data of the Dalls Preschool Test.

Barbara Kay Mann and Wendy Sherman have made a contribution worth sharing with others who work with very young children and Florida Educational Research and Development Council, Inc. is pleased to desseminate these findings for our readers.

CHARLIE T. COUNCIL, Executive Secretary

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION 1

Problem and Its Context 1

Research Hypotheses 2

CHAPTER II

REVIEW OF THE LITERATURE 3

General Issues in Preschool Screening 3

Psychometric Issues in Preschool Screening 5

CHAPTER III

METHOD 11

Study Design 11

Subjects 11

Instrumentation 12

Data Management 15

CHAPTER IV

RESULTS 17

Correlations Between DPST and CTBS 17

Multiple Regression with CTBS Reading Total as Dependent Variable 19

Multiple Regression with CTBS Math as Dependent Variable 20

Additional Analysis 20

CHAPTER V

DISCUSSION AND RECOMMENDATIONS 22

Pearson Correlations and Mulitple Regression 22

Considerations of Confounding Variables 24

Limitation of Instruments 25

REFERENCES 28

FLORIDA EDUCATIONAL RESEARCH AND DEVELOPMENT COUNCIL, INC.

OFFICERS:

President, Phyllis NeSmith
Treasurer, Don Stone

BOARD OF DIRECTORS:

Lee Rowell, John Hilderbrand, Evelyn J. Lynn, Jackson McAfee, Rick Nations, Raymond Turner, Mary Giella, Terrance Horrigan, Fran Kochan, James Daniel

ADVISORS TO THE BOARD:

Charles Hayes, Citrus County School Board
Charles Henderson, University of Florida
Cecil Carlton, State Department of Education
Sandra K. McDonald, University of North Florida
Jake Beard, Florida State University
Patricia Manning, University of Central Florida
Sharon Kossack, Florida International University
Hayman Kite, Florida Atlantic University
William Castine, Florida A & M
Sam Matthews, University of West Florida
John Follman, University of South Florida

EXECUTIVE SECRETARY:

Charlie T. Council, Florida Atlantic University and Southwest Florida Teacher Educational Center

MEMBERS COUNTIES	SUPERINTENDENTS	COUNCIL MEMBERS
Alachua County	Dr. Douglas Magann	Dr. Douglas Magann
Charlotte County	Robert L. Bedford	John Weigman
Citrus County	Roger Weaver	Carl Austin
Collier County	Dr. Thomas L. Richey	Donald Thomsen
Dade County	Dr. Leonard Britton	Ray Turner
De Soto County	Larry Browning	Phyllis NeSmith
Dixie County	James E. Hurst	Lloyd Jones
Glades County	Lester Mensch	Lester Mensch
Hamilton County	Fred D. Greene	James Coe
Hardee County	John E. Terrell	John Maddox
Hendry County	Dr. William Burke	Ken Dooley
Hernando County	Dr. James K. Austin	Martin Youngmann
Highlands County	Dr. Billy C. Cason	David W. Robinson, Jr.
Hillsborough County	Dr. Raymond Shelton	John Hilderbrand
Indian River County	Dr. James A. Burns	Jackson McAfee
Lee County	Dr. James E. Melvin	Donald Stone
Madison County	Randall M. Buchanan	Bill Roberts
Marion County	H. Leon Rogers	Lee Rowell
Martin County	V. James Navitsky	Terrance Horrigan
Nassau County	Craig Marsh	James T. Northey
Okeechobee County	Danny Mullins	Danny Mullins
Palm Beach County	Thomas Mills	James Daniels
Pasco County	Thomas E. Weightman	Mary Giella
Sarasota County	Dr. James Fox	Rick Nations
Suwanne County	Frank Strankunas	Marvin Jones
Taylor County	Glenda Hamby	Merle F. Knowles
Union County	James Cason, III	Howard G. McNeil
Volusia County	James E. Surratt	Evelyn J. Lynn
Wakulla County	Dr. William E. Whaley	Fran Kochan

CHAPTER I
INTRODUCTION

Problem and Its Context

There has been much concern in recent years with the identification of children with learning problems early in the educational process. (Segerstrom, 1976; Klein, 1978; Mardell and Goldberg, 1976). The passage of the Economic Opportunity Act of 1964 and the Elementary and Secondary Act of 1965 led to the development of assessment devices and programs that would measure and increase intellectual and social competence prior to school entrance. (Kresh and Green, 1969). New attention has been focused on the screening process with the 1975 passage of Public Law 94-142. Many school systems presently screen children before entrance to kindergarten using various instruments (Keogh and Becker, 1973). Screening results are typically used to provide information concerning a child's educational needs so that teachers can individualize learning (Casey, 1976). Individualization, whether in the regular classroom or in a special program, is designed to prevent or minimize development of other problems (Keogh & Becker, 1973). Because of this attention, focus has been placed on the Psychometric qualities of preschool screening devices, many of which lack adequate norms and reliability information. Even for those instruments with such information psychometric qualities are typically poor to moderate (Silver, 1976) (Silver and Hagin, (1981).

The purpose of this study is to provide information on the Dallas Preschool Screening Test (DPST). While the DPST is administratively feasible and is widely used in Texas, a paucity of research exists on the instrument. The current study provided data on the predictive validity of the DPST to total and component scores of the Comprehensive Test of Basic Skills (CTBS).

Research Hypotheses

The hypotheses under investigation in the present study were twofold:

1. There will be a significant correlation between the total and/or partial scores of the DPST and the total and/or partial scores of the CTBS.

2. There is a subset of the components of preschool performance as measured by the DPST which is better than any single predictor of kindergarten reading performance as measured by the CTBS.

CHAPTER II
REVIEW OF THE LITERATURE

The review of literature related to the present study will be presented as follows:

1. General issues in preschool screening

2. Psychometric issues in preschool screening

General Issues in Preschool Screening

A review of the literature reveals the direct goal of preschool screening to be the identification of children who are likely to develop learning problems. This identification can then lead to further assessment and resolution of the problem prior to the first grade (Reynolds, 1979). Preschool assessment offers the chance of detecting problems before they become so severe that remediation is more difficult (Gordon, 1974), (Wendt, 1978). Gordon (1974) notes that, for maximum success, intervention in learning problems should occur during the crucial early childhood period. Wendt (1978) contends that in order to be useful at all, screening devices should be related to developmental process rather than rote recall. He states that in order to make sound decisions, it is necessary to know those problems identified by both norm-based and criterion-referenced screenings and the type of kindergarten experiences that will be provided for the child.

There are several aspects of preschool screening which mitigate its usefulness. As intervention is typically a result of preschool screening, the two will be discussed together for the purposes of the discussion.

Harlow and Suomi (1970) raised the concern that early screening and possible resultant intervention typically requires young children to be separated from their mothers for too much of the day, thus affecting the children adversely. VanLeeuwen and Pomer (1969) reported similar conclusions in a study of separation adaptation in children aged two to five. In their

summary they state that circumstances around the separation had an impact on the amount of separation anxiety in a child and that there is a significant impact, particularly if it occurs concurrently with other traumatic events. Conflicting conclusions were reached by Caldwell, Wright, Honig, and Tannebaum (1970) who explored mother-child attachment in a group of home-reared versus day-care children. No weakening of the mother-child relationship was found in the day-care group as compared with the home-reared group.

A second issue deals with the charge that preschool screening promotes the labeling of children (Wendt, 1978). The evidence for and against labeling is complex and inconclusive. It is recognized that labeling does not necessarily occur for intervention to take place. However, since labeling can and often does occur, cases for and against labeling are presented.

Arguments made against labeling are that it is irreversible, that it may become a self-fulfilling prophecy, and that it is detrimental to the self-concept. Hobbs (1975) and MacMillan (1982) concur that children labeled and placed in special classes seldom return to regular classrooms. Possible reasons for this tendency to retain children in special programs beyond the point of efficacy include difficult separation between teachers and children and extension of goals beyond the criteria for effective functioning (Hobbs, 1975). Erickson (1964) too describes the labeling ceremonies in our culture as irreversible and suggests that labeling sets into motion a self-fulfilling prophecy. This prophecy claims that the child learns and accepts the applied label and proceeds to act in accordance with it, thus confirming the diagnosis (Guskin Bartel, and MacMillan, 1975). A related issue is that of negative effects of self-concept. According to this hypothesis the labeled child is treated differently by society and comes to believe that she is a devalued person (MacMillan, 1982).

There are several justifications for labeling. Hobbs (1975) states that labeling can open needed opportunities for children, aid in appropriate program funding, and provide a focus for local, national and world interest and support. The World Health Organization (1967) supports labeling by indicating that is is an

important aspect of research in any area. A major justification for labeling is the 1975 enactment of Public Law 94-142 which requires school districts to locate, diagnose and label exceptional children. Without the label children cannot be served in special education programs.

As noted, there are legitimate concerns relating to the possible results of screening preschool children. While these concerns need to be addressed, the benefits of preschool screening appear sufficiently valuable to warrant continuation of screening as well as further research in the area (Neisworth, Willoughby-Herb, Bagnato, Cartwright and Laub; 1980), (Wallace and Larsen, 1978).

Psychometric Issues in Preschool Screening

As an aim of screening the preschool child to obtain an accurate picture of performance, the issue of psychometric qualities of preschool screening devices arises. Four potential psychometric problems that confound the screening process will be discussed, test norms, reliability, content validity, and predictive validity. These psychometric issues relate to general and pre-school screening. Problems unique to pre-school screening are presented within each area.

Test Norms. Norms are used as the main frame of reference for interpreting and comparing a child's score to a specific group representative of the child (Gardner, 1962). One problem arises when information as to the norming sample is not provided, as is the case in many preschool screens (Silver, 1976). Eighty percent of the preschool screening instruments reviewed in Buros (1978) either lack reports of normative samples or provide inadequate information.

A second norming problem is the number of preschool screens which report normative information on the basis of small groups, frequently less than 100 per age level. Often these small groups are claimed to represent children from a variety of backgrounds (Boehm and Sandberg, 1982: Reynolds and Gutkin 1982). Though this claim is suspect, only sometimes do test authors urge caution in interpretation.

The DPST was normed with 3,000 three to six year old children in the Dallas suburb of Richardson, Texas. The children were from families described as above average in educational, social and financial status. At least half the population was enrolled in private schools and approximately one hundred black children were included. The test author noted that the norming population was different from a "global" population.

Reliability. The reliability of diagnostic procedures always involves some amount of error. There is often a discrepancy between a child's measured capacity and his functional performance. They often have little interest in their performance, are difficult to coerce, and may be fearful or shy (Sattler, 1982), all of which can contribute to error.

Further error is sometimes due to the rapid and irregular growth patterns of children (Bloom, 1964; Hunt, 1961). Children grow, change and experience new things from day to day, and their behavior is likely to change from situation to situation. This changing behavior of young children is unstable and leads to lower estimates of reliability than are found with older children (Reynolds and Gutkin, 1982).

Perceival (1972) reported that one class of 23 children was used to demonstrate reliability of the DPST. He makes no mention of the age group. Test-retest with a two-week interval was used. Pearson Product Moment correlation coefficients reported for each of the five subtests were Psychological,. 77; Auditory,. 84; Visual,. 95; Language,. 76; and Motor,. 84.

Lambert (1980) investigated the reliability of the DPST using a sample of 51 kindergarten children attending The Florida State University Developmental Research School in Tallahassee, Florida, in the school year 1979-1980. The children represented the school-age population of the State of Florida in terms of race, sex, socio-economic status, and academic ability. The DPST was administered and readministered with a 2-week interval. For the total DPST score, the test-retest reliability coefficient (r . 81) was statistically significant at the. 001 level.

Content Validity. Content validity is of importance in preschool screening; the test must stand by itself as an adequate measure of what it is supposed to measure. Frequently preschool screens cover multiple age levels and include few items appropriate for preschool children (Boehm and Sandberg, 1982). Additionally, it is not usually clear if a preschool screen is a representative sample of the universe domain as few of them have been subjected to rigorous content validation study (Barnes, 1982), (Lambert, 1980. Although some authors consider a sample which is not representative of the universe domain appropriate for screening purposes, many test consumers inappropriately translate possible hypotheses into "findings" (Reynolds and Gutkin, 1982).

A further problem of content validation is the many screening measures that have been constructed using items from developmental scales (Barnes, 1982). While there is nothing "wrong" with this procedure, most developmental scales were not constructed for screening and their relevance to identifying children with learning problems is uncertain.

The DPST is designed to screen six primary learning areas of 3 to 6 year old children. Further discussion of the six content areas is included in Chapter III. While the DPST appears valid, its actual effectiveness in measuring psychological, auditory, visual, language and motor areas is not known (Lambert, 1980).

TABLE 1

Test-Retest Correlation Coefficients for DPST

DPST subtests	(N = 51) r Lambert (1980)	(N = 23) r Perceival (1972)
Psychological	.73	.77
Auditory	.80	.84
Visual	.53	.95
Language	.79	.76
Motor	.76	.84
DPST total	.81	- -

Predictive Validity. A basic issue, and the focus of the present study, is the problem of the predictive validity of preschool screening devices which is typically uncertain or poor (Keogh and Becker, 1973; Silver, 1976). This is a result of several general problems in ensuring predictive validity.

The question of selection of examinees for entrance to a particular program is a frequent issue. (Harris, 1982). The problem relates to whether or not a low or high scorer on the predictor is permitted to enter the program. Ideally, all examinees should be admitted, however frequently lower high scores are excluded from programs due to economic or space reasons. If low or high scores are excluded, the variance of scores is much smaller and has a more restricted range than it would have if all scores had been considered. This results in lower measured validity. In the present study there is no exclusion of examinees on the basis of predictor scores.

A second threat to predictive validity is criterion contamination. this situation occurs when individuals gathering criterion data are aware of predictor scores allowing for judgemental influence on the criterion scores. (Harris, 1982). In the present study the criterion data gatherers, the kindergarten teachers, were unaware that the study was being conducted and did not have access to the predictor scores.

In order for criterion related validity to be meaningful, it is necessary for the selected criterion measure to be adequate, (Harris, 1982) Thorndike and Hagen (1977) specify the following desirable characteristics of a criterion measure: relevance, freedom from bias, reliability and availability. Review of the instrumentation section of the current study reveals that the CTBS meets the aforementioned characteristics.

Single and Multiple Variable Predictors. Several studies have been conducted to investigate the predictive validity of various preschool screening tests. Jansky and deHirsch (1972) named two categories of preschool predictors: single variable predictors which measure one broad area (e.g. visual motor integration) and multiple variable predictors which measure a combination of variables (e.g. test batteries). The following is a discussion of the predictive validity of tests in each of these categories.

Many single variables have been used as measures for prediction of achievement. Visual motor integration has been examined with the Bender Visual Motor Gestalt Test (BVMGT) (Bender, 1938. The BVMGT is frequently used for preschool screening and was investigated by Keogh (1965) who found correlations between the BVMGT and reading criteria to be statistically significant but not of significant magnitude for individual prediction of reading. Norfleet (1973) concurred with Keogh in finding the BVMGT weak in identifying poor readers (r -. 30). Keogh and Smith (1967) conducted a follow-up study to the Keogh (1965) study and again found relationships between the BVMGT given in kindergarten and grade 3 achievement to be low and non-significant (correlation coefficient not given).

Language has also been studied as a single variable predictor. Haring and Ridgeway (1967) studied thirty-one variables as predictors of learning problems. While they found language related variables most significant, accounting for 20 percent of the commonality in the principal components analysis, they inferred that when used alone, each of the variables had little use as a predictor. The predictive facility of intelligence test scores has been examined by Cohen (1963) and Segerstrom (1976). In a study of 56 kindergarten children, Cohen (1963) attempted to determine whether or not the Wechsler Intelligence Scale for Children (WISC) IQ was predictive of Metropolitan Achievement Test scores. Results indicated low predictive validity between the WISC IQ and Metropolitan Achievement Test scores (r . 35). Segerstrom (1976) investigated the predictive validity of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) Full Scale score to four criterion variables: Personal Social Behavior Rating Scale, Grades, Achievement Test scores and Passing or Failing a grade level. The WPPSI best predicted achievement test scores (r . 65) however poorly predicted whether the child passed or failed the grade.

The evidence clearly indicates that single variable assessment devices are not sufficiently dependable as predictors of achievement for preschool children. Use of multiple variable predictors would seem more reliable and logical for predictive purposes, as the use of more than one predictor would tap a broader experience base.

Jansky and deHirsch (1972) in a major study of multiple variable predictors derived a screening index of five tests from a battery of nineteen. When administered in the Spring of the kindergarten year, this index correctly identified 77 percent of the children who failed reading at the end of second grade. The multiple correlation coefficient of relationship between these five index tests and the end-of-second-grade silent paragraph reading achievement was. 69.

Segerstrom (1976) studied the relationship between forty predictor variables and eight criterion variables and found the majority of them to have no predictive ability. Findings were that those tests relating most highly to indicators of success emphasized both sensory motor and cognitive development.

The Denver Developmental Screening Test (DDST) (Frankenburg and Dodds, 1970) assesses children ages 2 weeks through 6 years in four areas: Personal-Social, Fine Motor Adaptive, Language, Gross Motor. Frankenburg and Dodds, authors of the DDST, report only concurrent criterion related validity for the DDST. Another major drawback is the normative sample which consisted of 1036 non-random, non-stratified Denver Children. This sample requires extreme caution in generalizing to other populations.

Authors of the Developmental Indicators for Assessment of Learning (DIAL) (Mardell and Goldenburg, 1975) report predictive validity of the DIAL to a combination of subtests from the Iowa Test of Basic Skills, the Metropolitan Achievernent Test, the Stanford Achievement Test, and the Metropolitan Reading Readiness Test. Reported correlations were highest for low SES children (r . 73) and. 65 and below for all other groupings of children: male/female, black/nonblack, urban/nonurban, nonlow SES.

As the above evidence strongly suggests, multiple variable predictors should be used with caution in predicting school success.

The lack of information on the predictive validity of the DPST indicates a need for investigation of the issues addressed in the current study.

CHAPTER III
METHOD

Study Design

An ex post facto correlation design was used in the present study. As the attributes measured by the DPST and the CTBS are fixed and not assignable, the purpose of this study was to test hypotheses that deal with the degree of relationship between the variables and does not assume cause and effect interpretation.

Subjects

The Center for Children with Special Learning Needs is a project with the purpose of identifying children with special learning needs and providing intervention service for those children enrolled in private preschools in Hillsborough County. In the Fall of 1981, the Center screened the population of children in the current study using the DPST. For those children showing developmental delay on the basis of the DPST scores, further assessment was conducted and appropriate intervention provided.

Many of these children tested by the Center have since enrolled in and are attending Hillsborough County Public School kindergartens where they were evaluated on the CTBS in April of 1983. These data were obtained through Hillsborough County Public School records.

The population of this study consists of those children with both Fall, 1981, DPST scores from the Center and Spring, 1983, CTBS scores from Hillsborough County Public Schools, an urban district with a population of 646,960 as of the 1980 census. The children ranged in age, at the time of DPST testing, from 4 years, no months to 4 years, 11 months. There were 53 females and 54 males; 12 were black, 92 were white, and 3 were Hispanic. The total population consisted of 107 children.

Selection of these children for DPST testing was a function of the voluntary participation of their school or day care center in the Center's screening program. Therefore, the selection was not random and generalization to other groups requires caution.

Instrumentation

The Dallas Preschool Screening Test. The Dallas Preschool Screening Test (DPST) was developed as an instrument for screening strengths and weaknesses in learning areas of preschool children (Perceival, 1972). It is designed to screen six primary learning areas for 3 to 6-year-old children and is scored on a successful/not successful basis. The DPST requires 15-20 minutes for administration and can be administered by the classroom teacher.

The DPST provides scores for six developmental areas. The Psychological subtest measures vocabulary and number concepts, receptive and expressive communication. The Auditory subtest is comprised of items measuring fine discrimination, perceptiveness, and memory. Included in the Visual subtest are paper and pencil tasks, coloring and color discrimination, and matching geometric designs. The Language subtest requires the identification of body parts and subjective scoring of child's verbal responses to pictures. The Motor subtest measures gross and fine motor assessment.

A Total score is also computed based on the combination of these five subtests.

Approximately 3,000 children were involved in standardization of the DPST. The technical report indicates that the population was about evenly divided into 3, 4 and 5-year-olds and that random samples were drawn. There is inadequate information as to whether the 3,000 children were the sample or were the population from which samples were drawn. The school district used was Richardson, a suburb of Dallas, described as above average in educational, social, and financial status.

At least half the population tested was enrolled in private schools and approximately one hundred black children were included in the population. It was reported that there was no significant difference in means between either the private school group or the black group and the total sample group. While the authors conclude that this information is sufficient to warrant the use of the DPST as a culture-free instrument, equal group means do not assure an unbiased test, the standard deviation for each group must be considered.

As reported by the author in the Manual (Perceival, 1972) DPST scores for thirty children from the 5 year age group and thirty children from the 4 year age group were compared to the Columbia Mental Maturity Test in a validity study. Pearson Product Moment correlation coefficient for the 5-year-olds was. 68 and for the 4-year-olds was. 59.

Additional evidence for the validity of the DPST was obtained by Perceival (1972), and the following results were reported.

Language DPST and Peabody Picture Voc. Test	r = .46
Motor DPST and Detroit Motor Sub. Test	r = .56
Psy. DPST and Draw a Man (Goodenough)	r = .59
Totai DPST and Draw a Man (Goodenough)	r = .55

While the psychometric qualities of the DPST have not yet been adequately displayed, in view of the previously discussed research by Lambert (1980) and Perceival (1972), as well as the administrative feasibility, it appears to have good potential for use as a screening instrument.

The Comprehensive Test of Basic Skills, Forms U and V

Final reliability and validity information on the CTBS form U and V were not available at the time of the current study. The following information was reported in the Comprehensive Test of Basic Skills, Form U and V, Preliminary Technical Report (1982).

The CTBS is a series of norm-referenced tests for kindergarten through 12th grades. Its purpose is to measure achievement in basic skills typically included in state and district curricula. These skill areas are Visual Recognition, Sound Recognition, Vocabulary, Comprehension, Total Reading, Math Concepts and Application.

Forms U and of the CTBS were standardized in the Fall of 1980 and Spring of 1981 with a large national sample of children in grades kindergarten through twelve.

Stratification by geographic region, community type (urban, suburban, rural), district size (average elementary enrollment), and a demographic index based on community characteristics related to district achievement was used for the public school sample. Stratification for the Catholic School sample was limited by region and district size and all other districts included in the sample were stratified by regions.

Content categories were determined by examination of state and district curriculum guides, texts, instructional programs and criterion-referenced assessment instruments. The process classifications were based on Bloom's taxonomy and includes recall, explicit information skills, inferential reasoning, and evaluation.

The visual recognition subtest consists of 18 items that measure a student's ability to distinguish upper and lower case letters. Tasks include recognition of orally presented single letters, recognition of upper and lower case of the same letter, and matching of letter groups.

The Sound Recognition subtest is comprised of 17 items that measure a student's ability to recognize sounds in spoken words. Tasks include identification of word pairs as being the same or different, identification of pictured objects that lower the same initial consonant sound, and identification of a pictured object that rhymes with a given word.

The Oral Comprehension subtest incorporates 15 items that measures understanding of spoken sentences and short

passages. In some of the items, the student identifies the picture that shows the meaning of a noun, verb, or preposition in a sentence. 0ther items require the student to identify the picture that illustrates details or ideas given in a passage.

The Vocabulary subtest includes 16 items that measure a student's ability to recognize categories and definitions.

These items require the student to identify the picture or word that is associated with an oral definition or category.

The Math Concepts and Applications subtest consists of 15 items that measure the student's understanding of mathematics concepts. Specific skills include concepts dealing with numeration, geometry and measurement units.

There are areas covered by both the DPST and the CTBS which are comparable in content. Both measure vocabulary and number concepts, memory, visual and auditory discrimination and receptive communication. These similarities provide a logical relationship between the two instruments and make the CTBS a particularly appropriate criterion for the DPST.

Data Management

Data Analysis. In order to test the predictive value of the DPST, three levels of data analysis techniques were employed.

Initially, an analysis was conducted using Pearson product moment correlations to test the hypothesis that there would be a significant (p . 05) correlation between the partial and/or total scores of the DPST and the partial and/or total scores of the CTBS. A 12 X 12 matrix was constructed using the six scores from the DPST and the six scores from the CTBS. Pearson Product Moment correlation coefficients were calculated for each matrix cell. Pearson Product Moment correlational analysis calculates the degree, the direction, and the significance of the relationship between the two sets of data. (Devoid, 1976). The assumption requirements of the Pearson Product-Moment coefficient are the following. First, there must

be a linear relationship, a straight line must describe fairly well the trend, regardless of how much above and below the line points may scatter. Second, each of the variables must be normally distributed. Third, the relationship must be homoscedastic, the spread about the best fitting line must be approximately the same at all levels of the two variables. (Nunnally, 1978). For the purpose of these analyses it has been assumed that the assumptions have been met. However, the possibility exists that the DPST and the CTBS are not linearly related and that their relationship is more complex. It is noted that visual examination of the scattergram did not indicate a more complex relationship.

The second level of analysis, multiple regression, was performed to examine the relationship of the DPST subtest scores to the criterion, the Total Reading CTBS scores. The hypothesis tested was that there is a subset of the component of preschool performance as measured by the DPST which best predicts kindergarten reading performance as measured by the CTBS. Reading achievement is frequently used as a criterion for school success (Keogh and Smith, 1967), (Segerstrom, 1976), (Jansky and deHirsch, 1972). Multiple regression analysis assumes that the battery of predictors measures two or more unrelated factors underlying the behavior to be predicted and that these subtest factors are related to the criterion behavior in such a manner that any deficiency in one ability can be exacerbated or compensated for by a sufficiently low or high level of performance in another ability (Harris, 1960). Multiple regression requires that variables are measured on at least an interval scale, a requirement which the DPST and the CTBS appear to meet.

Third, additional analysis of the lowest six DPST scorers and the lowest seven CTBS scorers was performed to determine if extreme scores reflected a better linear relationship than was represented by the entire sample.

CHAPTER IV
RESULTS

Pearson product moment correlation and multiple regression analysis were used to assess the degree of correlation between the DPST and the CTBS and to test the predictive validity of the DPST with the CTBS Reading Total score. The research design utilized two hypotheses which were directional as a basis for testing. Directional hypotheses specify whether the anticipated relationship will be positive or negative in nature. Consistent with the directional hypotheses, one-tailed tests of significance were utilized throughout the analysis. The .05 level of significance was used as the cutoff point for determining significance. Use of the .05 level infers that an outcome would occur only five percent of the time or less by chance.

The results of the study are discussed under four headings: degree of correlation between DPST subtest and total scores and CTBS subtest and Total scores, multiple regression with CTBS Reading Total as dependent variable, multiple regression with CTBS Math as dependent variable, and informal analysis.

Correlations Between DPST and CTBS

Pearson product moment correlations were calculated for subtest and total scores of the DPST and the CTBS. The Pearson r serves a dual purpose. Besides being an indicator of the goodness of fit of the linear regression of the DPST and the CTBS, it is a measure of association indicating the strength of the linear relationship between the two variables. Each of the DPST scores was examined for strength and direction with each of the CTBS scores. Correlation coefficients are presented in Table 2.

The DPST Visual scores correlated ($r = .26$) significantly at the .01 level with the CTBS Reading Total scores. It is noted that while this correlation is statistically significant it is not large enough to be predictive. No other DPST subtests correlation significantly with the CTBS Reading Total.

Five of the six DPST subtests correlated significantly with the CTBS Math subtest scores at the .05 level or better. The correlation between the DPST Auditory subtest and the CTBS

TABLE 2

CORRELATION COEFFICIENTS FOR DPST AND CTBS

CTBS	DPST PSYCHOLOGICAL	DPST AUDITORY	DPST VISUAL	DPST LANGUAGE	DPST MOTOR	DPST TOTAL
VISUAL RECOGNITION	.234*	.160	.100	.241*	-.082	.196*
SOUND RECOGNITION	.164	.119	.204*	.209*	-.008	.191
VOCABULARY	.179	.171	.267*	.149	.046	.224*
ORAL COMPREHENSION	.072	.076	.189	.164	-.142	.111
READING TOTAL	.136	.135	.263*	.183	-.089	.185
MATH	.270*	.240*	.349*	.319*	.083	.343*

N = 107

* = significant at .05 level or beyond

One-tailed test of significance

Math subtest (r = .24) was significant at the .05 level while correlations between CTBS Math and the DPST Psychological (r = .27), Visual (r = .35), Language (r = .32), and Total (r = .34) were significant at the .01 level.

Each of the remaining DPST subtest scores correlated significantly with CTBS subtest scores, however the correlations were so low as to be of questionable practical value. The CTBS Vocabulary subtest correlated significantly with the DPST Visual (p = .01, r = .27), and the DPST Total (p = .05, r = .22) scores. The CTBS Sound Recognition scores correlated significantly with the DPST Visual (p = .05, r = .20) and the DPST Language (p = .05, r = .21) scores. The CTBS Visual Recognition scores correlated significantly with the DPST Psychological (p = .05, r = .23), DPST Language (p = .05, r = .24), and DPST Total (p = .05, r = .20) scores.

It was noted that the DPST Motor subtest correlated negatively, though not significantly with the CTBS Visual Recognition, South Recognition, 0ral Comprehension and Reading Total scores.

Multiple Regression with CTBS Reading Total as Dependent Variable

Multiple Regression analysis is utilized for the purpose of isolating a subset of available predictor variable that will yield an optimal predictor equation with as few terms as possible. In the current study a stepwise regression analysis was performed in which variables were entered only if they met certain statistical criteria. The initial variable was selected for entry on the basis of significance and strength of correlation. The order of inclusion for subsequent variables was determined by the respective contribution of each variable to explained variance, whether or not the correlation between the variable and CTBS Reading Total was significant.

In using the multiple regression procedure, with CTBS Reading Total as the dependent variable the DPST Visual subtest was entered first. This subtest alone accounts for 10.9 percent of the

variance. The second variable entered was the DPST Motor subtest which increased the multiple correlation from. 33 to. 34, indicating that 11.6 percent of the variance was accounted for by the DPST Visual and Motor subtests together. It is important to note that two of the items on the Motor subtest had very little variability in scores when administered to the four year olds in this sample. Therefore, the effect of the DPST motor subtest is derived primarily from only one item (number 15) which measured gross motor skills.

Multiple Regression with CTBS Math as Dependent Variable

As the CTBS Math scores correlated more strongly with the DPST subtest scores than did the Reading Total scores, it was decided to perform multiple regression using the CTBS Math scores as the dependent variable.

On the basis of statistical significance, strength of correlation, and contribution to explained variance, the DPST Visual subtest was entered first and the DPST Language subtest was entered second. The addition of this second variable increased the multiple correlation from. 35 to. 39, meaning that approximately 15 percent of the variability was accounted for by the DPST Visual and DPST Language subtests together. The third variable entered was the DPST Motor subtest; however, this addition rendered an insignificant change of. 002.

Additional Analysis

In order to determine if the extreme scores reflected better the linear relationship than was presented by the entire sample, a scattergram analysis of the lowest six DPST scorers and the lowest seven CTBS scorers was performed. This analysis did not reveal a linear relationship in any of the four quartiles.

The factors of sex, race, age and intervention were also investigated to determine common descriptors for each of the low-scoring groups. Of the lowest six DPST scorers three were female, three male; three were white, two Hispanic, one black; their ages spanned from 5 years, 6 months to 6 years, 3 months; the level of intervention each received was not notably different.

Of the lowest seven CTBS scorers, four were female, three male; four were white, one Hispanic, two black: their ages spanned from 5 years, 5 months, to 6 years, 1 month; the level of intervention received by each was not notably different. The investigation of sex, race, age and intervention factors for DPST and CTBS low scorers was inconclusive in determining factors by which low scorers might be identified.

CHAPTER V
DISCUSSION AND RECOMMENDATIONS

The purpose of this study was to provide information on the Dallas Preschool Screening Test (DPST). Specifically, the predictive validity of the DPST was assessed by investigating correlations with CTBS scores. It was hypothesized that the DPST Total and partial scores would significantly correlate with the CTBS Reading Total and that a subset of the DPST subtest scores would significantly correlate with the CTBS Reading Total score. Although results indicated some significant correlations the size of the correlations suggested that they have little predictive value.

Pearson Correlations and Multiple Regression

Correlations were significant between the CTBS Reading Total and DPST Visual ($r = .26$) and DPST Language ($r = .19$) subtests. Each of these DPST subtests has areas of content also represented in the CTBS Total. The DPST Visual and the CTBS Reading Total subtests both include visual matching tasks. The DPST Language subtest and the CTBS Reading Total subtest both measure aspects of vocabulary and picture identification. It is noted that the CTBS Reading Total contains a greater variety of items than do either the DPST Visual or DPST Language subtests which may account for the relatively low level of correlation.

The strength of correlation between the DPST and CTBS Math scores was greater than between the DPST and the CTBS Reading Total scores. However, these correlations were also too small to be of predictive value. With the exception of the DPST Motor subtest, all DPST subtest correlated significantly at the .05 level or better with the CTBS Math subtest. The correlation coefficient for the CTBS Math and DPST Language subtests was .32. The relationship between the item content of these two subtests is not readily apparent.

The DPST Total scores correlated with the CTBS Math scores .34. The inclusion of the DPST Motor subtest in the DPST Total score served to lower the CTBS math — DPST Total

correlation coefficient as it is the only DPST subtest which did not significantly correlate with the CTBS Math.

The DPST Visual subtest correlated most strongly with the CTBS Math subtest. Examination of the items on each of these subtests did not reveal common content. It is presumed that there is a common underlying factor measured by both.

The correlation coefficient between the CTBS Math and DPST Psychological subtests was .27. Examination of the DPST Psychological items reveals a sample of mathematical tasks such as counting and repeating digit spans which could account for the correlation of this subtest with the CTBS Math subtest. The correlation between the CTBS Math subtest and the DPST Auditory subtest ($r = .24$) explained the least amount of variance of any CTBS Math — DPST correlation. The existing relationship between the two subtests may be due to the nature of the DPST Auditory items, all of which involve an attentional factor important in Mathematics.

As previously reported the DPST Motor subtest correlated negatively with the CTBS Visual Recognition, Sound Recognition, Oral Comprehension and Reading Total subtests. Although taken individually none of these scores were significant, they were all in the same negative direction. Consideration was given to the cause of the negative correlations. A visual scan of the 107 individual scores on an item by item basis revealed that the major contributor to the variance in the motor subtest is the gross motor area. Possible lines of reasoning for these negative correlations include the lack of gross motor tasks in the CTBS, the possibility that children with good gross motor skills, are more active and less likely to attend to academics, or that the DPST Motor subtest measures an undetermined factor which is inversely related to the CTBS.

Results of the multiple regression analysis using the CTBS Math subtest as the dependent variable resulted in a higher level of explained variance (15 percent) than did the regression with CTBS Reading Total (11.6 percent) as dependent variable. Examination of DPST Visual and Language items and CTBS

Math items did not reveal content similarity which suggests as one possible explanation that an underlying but unknown common factor was measured by each.

Considerations of Confounding Variables

A major drawback in conducting ex post factor research was the increased probability of confounding variables which could affect gathered data (Kerlinger, 1979). When interpreting such research the influence of these variables must be considered.

One such uncontrolled variable is the approximately one and one-half year interval between the DPST testing and the CTBS testing, and the possibility of ensuing intervention. Intervention could have improved the skill level of those children scoring initially lower on the DPST and thus lowering the correlations between the DPST and the CTBS. To examine the question of intervention at the preschool level, the bottom six DPST scorers were examined for type of service provided by the Center. One child received a complete psychological evaluation and behavioral objectives were provided to the preschool teacher; that child scored within the average range on the CTBS Math subtest. The teachers and/or parents of two of the low scorers received consultation services only and both scored in the above average range on the CTBS Math subtest. Three of the six received no services other than the administration of the DPST. These three scored below the mean although within the average range. While the low scorers for whom the Center provided more than screening services did score above the mean, the other three scored within the average range and thus the Center's intervention for these six students does not explain their CTBS scores.

Another intervention might have occurred at the kindergarten level. Hillsborough County currently functions under the Primary Education Program (PREP) Bill (S.230.2312 FL Stat.), which requires the educational and health screening of all kindergarten students within the first eight weeks of their enrollment. The health screening involves evaluations of vision, hearing, height, weight, and social-emotional development. These skills are measured using the Santa Clara or the DPST. If

a student scores below his chronological age he can be referred for further testing. This testing is done by the educational diagnostician who makes recommendations for the student's educational planning. The recommendations are carried out by the kindergarten teacher. In order to determine the possible effect this intervention might have had on the current study, school records of the DPST low scorers were again examined for referral to and evaluation by the educational diagnostician. Results indicated that only one of the six was provided the services of the educational diagnostician, thus it would appear as though PREP kindergarten intervention did not significantly affect the results of the present study.

A second uncontrolled variable in this study is related to the selection of students tested on the DPST. All students tested were enrolled in private preschools and thus had some educational experience before and after DPST testing. The possibility exists that for children without preschool experience the DPST would better predict kindergarten CTBS Reading Total and/or Math subtest scores. Further research into this question might substantiate the hypothesis that the DPST works selectively on varying groups of children in predicting their kindergarten CTBS scores.

Limitation of Instruments

An additional explanation of the outcome of this study is the technical adequacy of the DPST. First, the DPST appears to have an insufficient ceiling which does not allow for appropriate discrimination between students. This occurs when there are not enough difficult items included in a test. The insufficient ceiling has the effect of restricting the range of scores which results in a deceptively low correlation between the variables (Runyan and Haber, 1980). In the present study the insufficient ceiling may have accounted for the low correlation between high DPST and high CTBS scorers, however it does not account for low correlation between low DPST and low CTBS scorers.

A second consideration of DPST limitations deals with the paucity of items on the screening device. The total test consists of fifteen items which are divided into five subtest areas. The highest number of items for any one subtest is four which might well be too small a sample upon which to base a subtest score, thus threatening content validity (Nunnally, 1978).

A third possible weakness of the DPST is that it may not measure factors that are indicative of kindergarten success as measured by the CTBS Reading Total score. In consideration of this question the areas covered by the CTBS and the DPST were examined. The CTBS Reading Total score is comprised of visual recognition of letters and letter groups, identification of like and unlike word pairs, recognition of initial consonant sounds, rhyming, understanding short sentences and passages, and recognition of categories and definitions by identification of pictures or words associated with oral definitions.

The DPST subtests similarly measure visual matching though not of letters, spoken sentence comprehension, and a type of vocabulary task wherein the student provides the name of an object by its function. However, DPST does not measure factors such as identification of letters, auditory discrimination, rhyming or passage comprehension. Examination of CTBS and DPST content resulted in a combination of shared and not shared content.

Beyond the issue of DPST limitations is the possibility that the cognitive structure for four year olds and six year olds is such that it is not possible to measure in a four year old the underlying factors that will predict their academic success at six years of age. Preschool children do not carry out systematic cognitive organization in the same manner as school aged children do. Preschoolers concepts of objects are preconcepts, and perceptions or sensory impressions rather than cognition or logic dominate these concepts (Tudor, 1981). This cognitive organization changes qualitatively in content and structure from one state of development to another (LaBarba, 1981). As the structure changes, the assimilation of new information occurs which allows for more organizational change and the development of a new or different cognitive structure.

The purpose of the current study was to provide information on the Dallas Preschool Screening Test (DPST). Specifically, the relationships between the DPST partial and/or Total scores and The Comprehensive Test of Basic Skills partial and/or Total Scores were examined for preductive validity. Although results of the study indicated some significant correlations, the size of the correlations suggested that they have little predictive value.

Two areas of recommendations will be presented based on the findings of this study. First, the issue of the use of the DPST as a screening instrument will be discussed. Because of the small size of the significant correlations, it is suggested that the DPST be used in conjunction with some other measure rather than by itself. For example, there is considerable evidence (Alberman, 1973; Barnes, 1982; Lichtenstein and Ireton, 1984) to suggest that successful achievement in school is influenced not only by specific abilities and acquired skills, but also by factors of socio-economic level of the family, extent of the mother's education, and cultural values of the parents.

A second area of recommendation considers future research. On the basis of the current study it is difficult to make a recommendation on whether or not preschool screening in general should be continued. However, it would seem that preschool screening should not be abandoned but that continued research in the area is indicated. Certainly further investigation into the technical adequacy of the DPST is warranted as the amount of information currently available is minimal. Additionally, the DPST might be examined and improved upon, for example in the areas of local norming and item content. Finally, exploration should be made into the aforementioned possibility of combining the DPST with another measure that would enhance its use as a predictor to the CTBS scores.

REFERENCES

Alberman, E. (1973). The early prediction of learning disorders. Developmental Medicine Child Neurology, 15, 202-204.

Anastasia, A. (1976). *Psychological testing, fourth edition.* N.Y.: Macmillan.

Barnes, K.E. (1982). *Preschool screening. The measurement and prediction of children at risk.* Springfield, Ill.: Charles C. Thomas.

Bender, L. (1938). *A visual-motor gestalt test and its clinical use.* N.Y.: American Orthopsychiatric Association.

Bloom, B.S. (1964). *Stability and change in human character.* New York: Wiley.

Boehm, A.E. and Sanberg, B.R. (1977). Assessment of the pre-school child. In C.R. Reynolds and T.B. Gutkin (Eds.) *The handbook of school psychology.* N.Y.: Wiley.

Buros, O.K. (Ed) (1978). *Eight mental measurements yearbook.* Highland Park, New Jersey: Gryphon Press.

Caldwell, B., Wright, C., Honig, A. and Tannebaum, G. (1970). Infant day care and attachment. *American Journal of Orthopsychiatry, 40,* 397-412.

Cohen, T.B. (1963). Prediction of underachievement in kindergarten children. *Archives of General Psychiatry, 9,* 444-458.

Comprehensive tests of basic skills, form U and V, preliminary technical report. (1982). Monterey, C.A.: CTB/McGraw Hill.

Cronbach, (1971). Test validation. In R.L. Thorndike (Ed) *Educational measurement.* Washington, D.C.: American Council on Education, 443-507.

Devoid, R.R. (1976). *The predictive validity of a multi-modal pre-school screening program.* Unpublished doctoral dissertation, University of North Carolina at Chapel Hill.

Encyclopedia of educational research, volume 4. (1982). American Educational Research Association.

Erickson, K.T. (1964). Notes on the sociology of deviance. In H.S. Becker (Ed), *The other side: perspectives on deviance.* Glencoe, Ill.: Free Press.

Fallen, N.H., and McGovern, J.E. (1978). *Young children with special needs.* Columbus: Charles E. Merrill.

Frankenburg, W.K. and Dobbs, J.B. (1967). The Denver developmental screening test. *Journal of Pediatrics, 71,* 181-191.

Gardner, E.F. (1962). Normative standard scores. *Educational and Psychological Measurement, 22,* 7-14.

Gordon, S. (1974). Compensating, remediating, innovating, and integrating: illusions of the poor. In G. Williams and S. Green (Eds), *Clinical child psychology: current practices and future perspectives.* New York: Behavioral Publications, 292-396.

Gudkin, S.L., Bartel, N.R., MacMillan, D.L. (1975). Perspective of the labeled child. In N. Hobbs (Ed) *Issues in the classification of children, vol. II.* Washington: Jossey Bass.

Haring, N. and Ridgeway, R. (1967). Early identification of children with learning disabilities, *Exceptional Children, 33,* 387-395.

Harlow, H.F. and Suomi, S.J. (1970). Nature of love - simplified. *American Psychologist, 25,* 161-168.

Harris, C.W. (Ed) (1982). *Encyclopedia of educational research.* New York: MacMillan.

Henryssen, S. (1971). Gathering, analyzing and using data on test items. In R.L. Thorndike (Ed) *Educational measurement.* Washington, D.C.: American Council on Education.

Hobbs, N. (1975). *The future of children.* Washington: Jossey Bass.

Hunt, J.M. (1961). *Intelligence and experience.* New York: Ronald.

Jansky, J. and deHirsch, K. (1972). *Preventing reading failure.* New York: Harper and Row.

Keogh, B.K. (1965). The bender gestalt as a predictive and diagnostic test of reading performance. *Journal of Consulting Psychology, 29*(1), 83-84.

Keogh, B.K. and Becker, L.D. (1973). Early detection of learning problems: questions, cautions and guidelines. *Exceptional Children,* September.

Keogh, B.K. and Smith, C. (1967). Visuo-motor ability for school prediction: a seven year study. *Perceptual and Motor Skills, 25,* 101-110.

Kerlinger, F.N. (1979). *Behavioral research: a conceptual approach.* New York: Holt, Rinehart and Winston.

Klein, A. (1978). The reliability and predictive validity of the slosson intelligence test for pre-kindergarten children. *Educational and Psychological Measurement, 38.*

Kresh, E. and Green, B.G. (1968, Feb.) Preschool academic skills test. Paper presented at American Educational Research Association Meeting, Los Angeles.

LaBarba, R.C. (1981). *Foundations of developmental psychology.* New York: Academic.

Lambert, D.M. (1980). The reliability and concurrent validity of the Dallas preschool screening test. *Dissertation Abstracts International. 71*(3), 930-A, 8020355.

Lichtenstein, R. and Ireton, H. (1984). Preschool screening identifying young children with developmental and educational problems. Orlando: Grune and Stratton.

MacMillan, D.L. (1982). *Mental retardation in school and society.* Boston: Little Brown.

Mardell, C.O. and Goldenburg, D.S. (1983). *Developmental indicators for the assessment of learning revised manual.* Edison, New Jersey: Childcraft.

Mardell, C.O. and Goldenburg, D.S. (1976, August). The predictive validation of a pre-kindergarten screening test. Paper presented at International Scientific Conference of IFLD, Montreal.

Neisworth, J.T., Willoughby-Herb, S.J., Bagnato, S.J., Cartwritht, C.A., Laub, K.W., (1980). *Individualized education for preschool exceptional children.* Germantown, Maryland: Aspen.

Norfleet, M. (1973). The bender-gestalt as a group screening instrument for first grade reading potential. *Journal of Learning Disabilities, 6,* 383-388.

Nunnally, J.C. (1978). *Psychometric theory.* New York: McGraw Hill.

Perceival, R.R. (1972). *The Dallas pre-school screening test - an assessment of learning pathways.* Richardson, Texas, Dallas Educational Services.

Primary Education Program: An act relating to education creating S230.2312 FL Statutes providing the primary education program.

Reynolds, C.R. (1979). Should we screen preschoolers? *Contemporary Educational Psychology, 4.*

Reynolds, C.R. and Gutkin, T.B. (Eds) (1982). *The handbook of school psychology.* New York: Wiley. Segerstrom, R.R. *A study in the prediction of school readiness and school failure.* Published Doctoral Dissertation, University of Northern Colorado.

Runyan, R.P. and Haber, A. (1980). *Fundamentals of behavioral statistics.* Reading, Massachusetts: Addison-Wesley.

Sattler, J.M. (1982) *Assessment of Children's Intelligence and Special Abilities,* Second Ed. Boston: Allyn and Bacon.

Silver, A. (1976). Basic handbook of child psychiatry. (incomplete information).

Silver, A. and Hagin, R. (1981) Bridge to tomorrow. In Silver and Cruickshank (Eds). *Selected papers from the 17th international conference of the association for children with learning disabilities. Syracuse: Syracuse University Press.*

Standards for educational and psychological tests. (1974). Washington, D.C.: Amercian Psychological Association. Thorndike, R.L. and Hagen, E.P. (1977). *Measurement and evaluation in psychology and education.* New York: Wiley.

Tudor, M. (1981). *Child Development.* New York: McGraw Hill.

VanLeeuwen, E.L. and Pomer, M.R. (1969). Separation anxiety in children age two to four. In C. Landreth, *Early childhood: behavior and learning.* New York: Knopf, 100-103.

Walizer, M.H. and Wienir, P.L. (1978). *Research methods and analysis, searching for relationships.* New York: Harper & Row.

Wallace, G. and Larsen, S.C. (1978). *Educational assessment of learning problems: testing for teaching.* Boston: Allyn and Bacon.

Wendt, R.N. (1978). Kindergarten entrance assessment. Is it worth the effort? *Psychology in the Schools, 15*(1).

World Health Organization. (1967). *International statistical classification of disease.* Geneva: World Health Organization.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION 34

Context of the Problem 34

Statement of the Problem 35

Purpose 36

CHAPTER II

LITERATURE REVIEW 37

Preschool screens - rationale and purpose 37

Single purpose vs. multi-purpose screens 39

Rating scale as criterion of kindergarten success 41

CHAPTER III

METHOD 45

Study design 45

Subjects 47

Instrumentation 47

Data management 50

CHAPTER IV

RESULTS 51

Predictive validity analysis 51

Multiple regression 52

Criterion comparison 52

Additional analysis 55

CHAPTER V

DISCUSSION AND RECOMMENDATIONS 56

REFERENCES 64

APPENDIX 1

HILLSBOROUGH EARLY RATING SCALE 70

APPENDIX 2

MEANS AND STANDARD DEVIATIONS OF DPST, HERS AND CTBS 73

CHAPTER I
INTRODUCTION

Context of the problem

Recent interest in early identification of potential learning problems and disabilities has fostered the development of numerous devices designed to screen the abilities and behavioral characteristics of young children (Attwell et al., 1967, Brigance, 1981, Conrad and Tobeissen, 1967, Kresh and Green 1969, Mardell and Goldenberg, 1983, Percival, 1972, Satz and Fletcher, 1982, Southworth et al., 1980). There is substantial evidence that early screening followed by appropriate intervention may affect a child's developmental potential and help to prevent later school failure (Bronfenbrenner, 1974, Hockman, 1972, Reynolds, 1979). Screening should occur as early as possible while there remains opportunity for the child to benefit from intervention and before any problems become cumulative (Meisels, 1976). Thus, the optimal time for screening a child appears to be prior to the child's entrance to kindergarten.

Finding the most effective technique for validly predicting learning related problems has been the subject of controversy (Bryan and McGrady, 1972, Keough and Smith, 1970, Wendt, 1978. Southworth et al. (1980) lists many instruments used in predicting elementary school success based upon kindergarten performance but fewer that are intended to predict kindergarten success from preschool performance.

Research supports the use of multi-purpose screening instruments for predicting school success (Casey, 1976, Lichtenstein and Ireton, 1984). For the purpose of this study, multi-purpose screens measure at least two types of behavior. They are given preference to single purpose screens because the multi-purpose instruments offer the examiner a wide range of information about the child.

Most preschool instruments, whether single or multi-purpose lack substantial norming reliability, and validity data (Rubin and Balow, 1968, Silver, 1977, Wendt, 1979). Without adequate

psychometric properties, even the use of wide ranged multi-purpose preschool screens can only be tenuous. Additional research is needed to establish a predictive relationship between instruments that measure preschool characteristics and those that measure later school success.

Statement of the problem

Consistent with the issues discussed in the previous section, this study was designed to contribute information about the predictive validity of one preschool screen: the Dallas Preschool Screening Test. This test compares favorably to other preschool screens in the adequacy of its technical properties (Percival, 1972) and has been used extensively in Hillsborough County. The relationship between preschool performance, as measured by the Dallas Preschool Screening Test (DPST), and kindergarten success, as measured by the Hillsborough Early Rating Scale was examined.

This study also examined the relationship between the Comprehensive Test of Basic Skills (CTBS) and the Hillsborough Early Rating Scale (HERS). Although the CTBS is a more traditional and more widely used criterion of academic achievement than the rating scale, the rating scale has many advantages. The HERS, by virtue of its being based on teacher observations of student behavior, provides a different, although not necessarily better, perspective of progress in school than the CTBS. The relationship between the CTBS and the HERS strengthens knowledge of the rating scale as a criterion of achievement and contributes to a more comprehensive understanding of the use of the DPST as a predictor.

A rating scale is used in this study as the primary criterion measure. Although research has supported the use of rating scales as predictors of future school success (Attwell et al., 1967, Glazzard, 1977, Virgin and Schecter, 1974) they have not customarily been used as criterion instruments. The factors that cause rating scales to be excellent predictors also suggest that they will be found to be strong criterion measures. Metham (1978) suggests that an important advantage of the rating scale is that teachers, who are equipped with developmental

information and who have observed the child over long periods of time, are effective raters. Concerns for such rating scale problems as ambiguity of items and teacher expectancy effects can be minimized by precise construction of items. Overall, the research has indicated that carefully constructed rating scales compared favorably with standardized tests as measures of early school success (Conrad and Tobeissen, 1964, Glazzard, 1982, Virgin, 1974).

Purpose

The purpose of this study was two-fold:

1. To investigate the predictive validity of the Dallas Preschool Screening Test.

2. To investigate the relationship between the Hillsborough Early Rating Scale and the Comprehensive Test of Basic Skills.

CHAPTER II
LITERATURE REVIEW

The evaluation of the literature of this study is organized into three sections.

1. Preschool screening rationale and purpose
2. Single-purpose vs. multi-purpose screens
3. Rating scale as a measurement of kindergarten success

Preschool screens—rationale and purpose

The purpose of any screening instrument is to provide a gross estimate of present developmental levels. It is not intended to make fine discriminations for diagnosis and placement into special programs. A screen identifies potential deficit areas and provides information to use as the basis for recommending further in-depth assessment.

Considerable controversy has surrounded the use of preschool screens. Research in this area can be divided between studies that explore the benefits of screening preschoolers and those that cite the dangers of the screening. The majority of the evidence is in support of preschool screening.

Four ideas are basic to the literature that mitigates against adequate screenings. First, Keough and Becker (1973), Reynolds (1979), and Wendt (1979) question the value of screening such young children because they are rapidly developing. Second, the effects of labeling children prior to their formal entrance to school may be harmful (Wendt, 1979). That label, regardless of its accuracy, will be carried with the children throughout their school careers. Third, Adelman (1971) and Wendt (1979) are opposed to preschool screening because the technical properties of the instruments being used are largely unacceptable. The predictive validity and reliability of such screens is highly questionable. Finally, Wendt (1979) also discusses the high cost in time, money and personnel that some screens require. Most of the authors cited who are opposed to

preschool screening are not opposed to early detection of problems but rather, question the ability of present techniques to accomplish the task.

The overwhelming basis for support of preschool screening is that it has been proven successful (Hockman, 1972; Lichtenstein and Ireton, 1984). The data indicate that early identification of learning problems with subsequent special programming has beneficial effects.

Authors such as Devoid (1976), Hinton and Knights (1971), McKnab (1972), and Segerstrom (1976), provide evidence that early identification can prevent later school failure. Unless problems are discovered early in the child's education, the evidence suggests that any learning problems will get more serious as the child grows older (Devoid, 1976).

Hinton and Knights (1971) studied 67 children with learning problems who had been referred to a psychologist for evaluation. The average grade placement at the point of referral was grade 4. The primary reason for referral was either the threat of retention, low grades, or disruptive behavior. The authors studied the academic history of the children and found that failure had occurred in almost all cases in grade 1. In two other studies the pattern of failure was similar. The authors concluded from this analysis that failure most often occurs early in the school career. Observations of early school failure by children who are later referred to a psychologist for learning problems indicate the importance of early detection and intervention. Undetected problems lead to continued academic difficulty and continued school failures (Hinton and Knights, 1971).

Keough and Smith (1970) and Haring and Ridgeway (1967) discuss the opportunity that early identification of potential problems affords kindergarten teachers for constructing appropriate individualized programs to meet the child's specific needs. Parents and teachers are becoming increasingly aware of the importance of early recognition of learning problems (Devoid, 1967, Haring and Ridgeway, 1967). Screening devices have the potential to be accurate sources of information about possible developmental deficiencies.

Single-purpose vs. multi-purpose screens

Existing preschool screening instruments may be divided into two groups: single-purpose and multi-purpose tests. Single-purpose tests measure one type of behavior such as receptive vocabulary, visual-motor integration, or motor skills. Multi-purpose tests measure two or more types of behavior and result in a wider range of information than a single purpose test.

Many screening attempts utilize the single-purpose test but most of these attempts produce only limited success in predicting learning problems. Rosenberg and Stroud (1966) used two single-purpose tests (Peabody Picture Vocabulary Test and Columbia Mental Maturity Scale) to predict potential achievement in a limited population of children. These tests over-estimated seriously the prevalence of retardation in 28 poverty-area children. The results indicate that neither of the two tests alone can accurately predict achievement for that population. Another study by Matheny (1971) compared PPVT scores to overall academic achievement for educable mentally handicapped children and found that the PPVT did not significantly correlate with academic achievement. These two studies substantiate each others findings that the PPVT is not an accurate predictor of early achievement.

Correlations between kindergarten scores on the Bender-Gestalt Test and later school achievement were not statistically significant in a study by Keough and Smith (1967). Koppitz et al. (1959) found that the Bender-Gestalt Test and Human Figure Drawings used together were useful school predictors but indicated that the prediction of school success should take into account as many factors as possible.

The results of these studies indicate that single-purpose tests are not comprehensive enough to adequately predict future school success. A child's learning is based on responses to a variety of environmental factors. For that reason, early prediction of learning problems must involve aspects of many factors. Daniels (1970), in a summary of his evaluation of investigations that attempt to predict school achievement, concludes that predictive validity will be greater if the

predictions are based on combinations of tests that measure different aspects of personality and intellectual capacity.

There are fewer multi-purpose screens than single-purpose screens but they offer greater promise for their ability to predict school performance. The Denver Developmental Screening Test (Camp, Frankenburg and Van Natta, 1971) is a short, easily administered test that measures four functions of preschool children: Physical, motor, perceptual, and cognitive development. The authors reported test-retest reliability as ranging from. 66 to. 93. A validity study found that the Denver correctly classified 60% of mentally deficient children.

Conrad and Tobeissen (1967) developed the Schenectady Kindergarten Rating Scales to screen kindergarten and pre-kindergarten children. The scales measure fourteen personality and achievement behaviors. The instrument was found to be an effective means of predicting first grade achievement in 63% of girls and 73% of boys.

The DIAL-R (Mardell and Goldenberg, 1983) measures the areas of gross motor, fine motor, concepts, and communications. The authors report total test-retest reliability for the DIAL-R as. 87 and predictive validity comparing the DIAL to a combination of subtests from standardized achievement tests as ranging from. 45 to. 73.

The Florida Kindergarten Screening Battery (Satz and Fletcher, 1982) combines pre-existing tests with new devices in a comprehensive screening that measures visual-perceptual, perceptual-motor, and soma-to-sensory abilities plus receptive vocabulary and alphabet recitation. The instrument reports reliability coefficients ranging from. 77 to. 98. A three year follow-up study reports predictive validity coefficients ranging from. 54 to. 91 using reading achievement as the criterion.

The Dallas Preschool Screening Test measures six areas of developmental readiness (Percival, 1972). Lambert (1980) used the Dallas Preschool Screening Test in an investigation to estimate the test-retest reliability, concurrent validity with the Metropolitan Readiness Test (MRT), and group performance

differences in age, sex, and conceptual background as measured by performance on the Tests of Basic Experiences (TOBE). The subjects were 51 kindergarten children who represented the school-age population of the State of Florida in terms of race, sex, socioeconomic status, and academic ability. The reliability was reported as. 81. The concurrent validity coefficient with the MRT was. 62. Both of these estimates were significant at the. 05 alpha level. Results of the study indicated that the DPST total score demonstrated acceptable reliability and validity. The research found low reliability for some subtests which was attributed to a lack of complete administration and scoring directions.

It appears that multi-purpose screens are superior to single purpose screens in predicting later school performance. A summary by Casey (1976) of instruments using this multidisciplinary approach recommends six factors to be included in the screening: motor development, visual-motor coordination, speech and language, social knowledge, auditory skills, and visual skills.

In summary, screening preschoolers for the purpose of detecting possible learning problems is an important challenge for many educators. Research indicates that multi-purpose tests have the potential to be effective predictors of school success. Many of these screening devices are lacking in adequate psychometric properties. Additional research in the development and validation of such instruments is required before preschool screens will become dependable as predictors of later performance.

Rating scale as criterion of kindergarten success

Readiness tests and rating scales are frequently used to predict early school achievement (Glazzard et al., 1982). Rating scales have traditionally not been the preferred measure of school progress and have often only been used as a backup instrument to the standardized test. Theoretically, the rating scale has a variety of characteristics which suggest that it may, in fact, be an important aspect of any evaluation of school progress.

An important advantage of the rating scale is that teachers, compared with other school personnel, have the closest, most continuous relationship with the child. Thus, they are better prepared to judge the child's behaviors than any testing specialist (Keough, 1973). The teacher is in possession of developmental knowledge about the child and is best able to recognize behaviors, abilities, and difficulties which have relevance to the school setting (Metham, 1978).

Bryan and McGrady (1972) and Glazzard et al. (1982) discuss a number of practical advantages of using teacher's rating scales. First, rating scales are economical in money and time. Traditional readiness tests are time consuming to administer and to score. Rating scales greatly minimize teacher time and eliminate student time (Glazzard et al., 1982). Second, the rating scale involves the teacher as the primary source of information. As mentioned above, the teacher bases his/her ratings on many hours of observations of the child's behavior (Bryan and McGrady, 1972). Third, rating scales often cover areas not usually testable such as peer relationships, oral language functions, and perceptual development. Tests are frequently only concerned with academic areas. Finally, an in-depth rating scale may identify specific areas of disability despite overall adequate performances (Bryan and McGrady, 1972). Often, a test yields only cluster or total scores which may mask a specific area.

Some of the possible negative aspects of rating scales such as teacher expectancies, the halo effect, and item ambiguity have been concerns for a variety of authors (Brophy and Good, 1974, Luce and Hoge, 1978, Metham, 1978). Metham cautions against rating scales that contain potentially ambiguous items. Behaviorally worded items that force a teacher to rate a specific observable behavior decrease ambiguity and create greater reliability in the scale. Behavioral items are thus, an important quality of any rating scale.

In a comprehensive review of the research in the area of the effects of teacher expectations on their rankings of student behavior, Brophy and Good (1974) found that teachers made stable and accurate assessments of a child's abilities and potential based on observations of that child. Teachers'

rankings of students depended more on the students' general self-control, school motivation, and learning abilities than on the presence or absence of particular knowledge. Ratings of student characteristics in non-academic areas are therefore, necessary in comprehensive rating scales.

Luce and Hoge (1978) also investigated the effects of teacher expectancy including the halo effect. They concluded that these effects were not operating as determinants of ratings of achievement.

In reviewing validity studies using rating scales, it was apparent that unlike the present study, rating scales have most often been used as the predictor of school success, not as the criterion. As predictors most rating scales were successful.

Glazzard et al. (1982) administered two measures to three districts of kindergarten students: a rating scale and a standardized reading readiness test. Results indicate that the rating scale had a significantly higher correlation with end-of-first-grade reading achievement (r . 53,. 67,. 57) than did the readiness test (r . 21,. 35,. 39). These authors concluded that the rating scale was more accurate in classifying students with or without reading problems than was the readiness test. The rating scale was able to correctly classify 59% of the students, the readiness test 49%, and the rating scale plus the readiness test 65% of the students. Although the rating scale was more accurate than the readiness test, both instruments together have the greatest accuracy in classifying students.

Virgin (1974) found a significant chi square between teachers' ratings and the word knowledge section of the Metropolitan Achievement Test and a significant chi square between teachers' ratings and the math of section of MAT. Thus, they concluded that kindergarten teachers are able to predict their pupils' subsequent performances. Conrad and Tobeissen (1967) were also successful in predicting first grade achievement for 63% of boys and 73% of girls with diagnostic profiles using a kindergarten rating scale. Attwell et al. (1967) found that behavioral ratings of kindergarten children significantly correlated with standardized achievement test results five years

later in five out of ten rated areas. Correlation coefficients ranged from. 26 to. 40 in these five areas.

All of these studies used a rating scale as the predictor. Although there is little research using rating scales as criterion measures, the attributes that make rating scales excellent predictors should also contribute to their effectiveness as criteria.

Studies like those of Attwell et al. (1967), Conrad and Tobeissen (1967), Glazzard et al. (1982) and Virgin and Schecter (1978) provide evidence in support of the use of rating scales as an alternative approach to standardized readiness tests in identifying students who may have achievement difficulties. Concerns for the negative impact of teacher expectancy and ambiguity have not been sustained by the works of Brophy and Good (1974), Luce and Hoge (1978) and Metham (1978). The practical implications of obtaining expert ratings of observed behavior, saving money and time, and obtaining ratings across many types of behavior contribute to the research findings that well developed rating scales may be as effective as standardized tests in predicting achievement.

CHAPTER III
METHOD

Study Design

This study was an ex-post facto correlational design to determine the predictive validity of the Dallas Preschool Screening Test using the Hillsborough Early Rating Scale as the criterion. Three statistical procedures were carried out in this study. Raw scores were used for all procedures.

1. a. Calculation of correlation between total DPST score and total HERS score.

   The null hypothesis states that there is no significant relationship between the total DPST score and the total HERS score.

   b. Visual scan for patterns of simple correlations in each cell of a 5 x 5 matrix (DPST subtests x HERS factors).

   The null hypothesis states that there is no significant relationship between each subtest of the DPST and any factor of the HERS.

2. Performance of a multiple regression to determine which subtest or which combination of subtests of the DPST best predicts the total HERS score.

   The null hypothesis states that there is no significant relationship between any series of subtests of the DPST and the total HERS score.

3. a. Calculation of the correlation between total HERS scores and total CTBS score.

   The null hypothesis states that there is no significant relationship between the total HERS score and the total CTBS score.

   b. Visual scan for patterns of simple correlations in each cell of a 5 x 5 matrix (CTBS subtests x HERS factors).

The null hypothesis states that there is no significant relationship between each subtest of the DPST and any factor of the HERS.

In this research design three specific rival hypotheses must be considered as threats to internal validity. First, any intervention following the screening could be an important confounding variable in a predictive validity design. Predictive validity assumes that the conditions existing at the start of the time sequence will still exist at the end of the time sequence. Any child who has received systematic intervention strategies during either preschool or kindergarten could potentially introduce error into the design. The DPST scores were not sent to the child's kindergarten teacher so that no direct intervention based on the DPST results could have occurred in the kindergarten year.

Second, there is question that the HERS may be an inappropriate criterion for the DPST. The HERS employs the child's class as the standard of comparison while DPST scores are compared to a normed population. In addition, the HERS is based on long-term observations of behaviors and the DPST is a measure of actual performances. Despite these differences between the two instruments, the HERS fulfills two basic requirements of criterion variables. A criterion should be chosen: a) in reference to the problem at hand and b) on the basis of beliefs about the nature of the ability reflected by the test scores (Davis, 1974). The HERS was chosen because it's factors appear to overlap with the DPST and the CTBS factors. In addition, it has the potential to distinguish between kindergarteners who are succeeding in school and those who are not (Barnard, Neville and Hilderbrand 1983).

Finally, total scores may be too general to indicate any useful relationships between predictor and criterion. This study included a visual scan for patterns of subtest correlations and a multiple regression analysis to study the effects of individual subtests and combinations of individual subtests.

Subjects

Approximately 482 children born between November 1, 1976 and November 1, 1977 were screened with the DPST in 1981-82 by staff members of the Children's Center at the University of South Florida. The children were enrolled in 26 private preschools and day-care centers in the Tampa, Florida area. Each school had applied for the services of the Children's Center so that the selection process was not random. All four year old children in each center were screened, not only those suspected of having learning problems. This study's sample was composed of those children from the original population of four year olds who were enrolled in Hillsborough County public school kindergarten classes in 1982-83 and who had end-of-kindergarten CTBS scores. The population consisted of 101 kindergarten children of which there were 51 males, 50 females, 88 whites, 11 blacks, and 2 hispanics. Certain limitations of this population considerably reduce the generalizability of the findings. Primarily, the sample did not take into account any possible differences between four year old children who were given the opportunity to attend preschool and those who were not. Potential characteristic differences between preschool programs operated by directors who requested the Children's Center services and those operated by directors who did not request services are also not considered. Finally, any of the original population who were enrolled in private kindergarten classes or any who did not go on to kindergarten were not included in this study.

In light of these limitations, the findings of this study may be generalized to those children in the population studied all of whom were enrolled in Hillsborough County public kindergarten classes in 1982-83 and attended one of the twenty-six preschools involved with the Children's Center in 1981-82. In addition, the results may have importance to other kindergarten children with preschool experience.

Instrumentation

The Dallas Preschool Screening Test. The Dallas Preschool Screening Test (Percival, 1972) screens six primary learning areas for children from three to six years of age. The learning

areas are: psychological, auditory, visual, motor, language, and articulation development. The test involves problem-solving tasks which are graded as either successful or unsuccessful as compared to the expected normal development of the child.

On a non-empirical level the DPST satisfies the requirements of preschool screens: brief, easy to administer, and inexpensive. The test takes approximately fifteen minutes to administer; it does not require extensive training so teachers can easily administer it; the manual and protocols are not expensive to purchase.

The DPST was standardized on approximately 3,000 three, four and five year old children in one suburban school district. Test-retest reliability with a two week interval was established. The reliability coefficients range from. 77 to. 95. Validity is reported as having been estimated with five different predictor/criterion pairs, but the level of validity was not reported. Validity coefficients ranged from. 46 to. 68.

The Hillsborough Early Rating Scale. The HERS was developed by the Department of Testing and Evaluation and the Kindergarten Diagnostic Team in Hillsborough County, Florida. The scale, in its present form, is composed of 32 objectively defined items. Each item is a five point scale on which the kindergarten teacher ranks the child in comparison to his or her classmates. The teacher rates the entire class as being of "high", "middle" or "low" ability or "some of each."

A study to investigate the psychometric properties of the Hillsborough Early Rating Scale has been undertaken at the University of South Florida, Tampa (Barnard, Neville and Hilderbrand, 1983). Two distinct populations of children were used: a public school population and a private school population. Factor analysis of ratings for both groups revealed almost identical factor loadings. The following five independent factors accounted for more than 75% of the variance: mental health, language development, cognitive functioning, motor development, and school achievement. The replication of the public school factor structure by the private school group indicates that the HERS demonstrates substantial reliability. Factor I accounted for the greatest amount of the total variance. Table 1 provides a breakdown of items in each factor.

The Barnard, Neville and Hilderbrand study (1983) also involved a validity analysis of the HERS. Two categories of children were identified: 1) those preschoolers who received only a DPST and 2) those preschoolers who received a DPST but also required a more in-depth evaluation. The children's kindergarten teachers completed the HERS one and one-half years after the DPST. The teachers had no knowledge of the initial screening. The study found a significant difference between the two categories of children on teacher ratings in the direction of positive ratings attributed to preschoolers who received DPST only on all five factors. In addition, a negative correlation between degree of services provided to the child as a preschooler and total HERS score administered in kindergarten was significant at beyond the. 01 level.

Table 1. Factor Structor of Rating Scale Items.

Factor	Name	Principle Items
I	Mental Health	14, 15, 16, 17, 19, 20, 21, 22, 27
II	Language Development	7, 8, 9, 19
III	Cognitive Functioning	1, 2, 3, 28, 31, 32
IV	Motor Development	11, 12, 13
V	School Achievement	26, 29, 30

The Comprehensive Test of Basic Skills. The CTBS (Form U) is a norm-referenced test for grades K-12 which was standardized in 1980-81 on a large national sample. The public school sample was stratified by geographic region, community size, school district size, and district achievement level. Reliability and validity data are included in the CTBS Technical Report which has not yet been published, but past test forms have yielded strong psychometric properties.

CTBS content areas were determined by state and district curriculum guidelines, texts, instructional programs, and criterion-referenced assessment instruments. Specific content

areas included in the CTBS (Form U) for kindergarten students are: visual recognition, sound recognition, vocabulary, oral communication, and math concepts. The CTBS was selected by Hillsborough County because its content matched the county's curriculum content.

Data Management

Data collection. The Children's Center at the University of South Florida used the DPST to screen 685 three, four, and five year old children between the Fall of 1981 and the Spring of 1982. Data for those 200 children in the population who were enrolled in public kindergarten programs in Hillsborough County in 1982-83 was collected through the Children's Center. Individual raw scores on each of the five DPST subtests as well as total raw scores were accessed from a computer file.

The HERS was sent by mail to the teachers of each of the 200 children. The rating scales were collected before the results of the CTBS were available to eliminate the possibility of CTBS scores biasing teacher opinions. Of the 200 rating scales requested, 181 were completed and returned.

Hillsborough County schools administered the CTBS (Form U) to all kindergarten children in the Spring of 1983. Permission to collect data from public school personnel was granted by Hillsborough County public schools. The county provided access to computer listings of all available raw scores. Only 101 of the original children had DPST, HERS, and CTBS scores.

Data Analysis. In order to examine the predictive validity of the DPST, two major statistical procedures were used: Pearson product-moment correlation matrix and multiple regression. One-tailed tests of significance were done throughout the analysis since the literature supports a positive relationship between preschool screening and kindergarten performance. The. 05 alpha level was the cutoff point for determining significance.

Pearson product-moment correlations determined the degree of relationship between total DPST scores and total HERS scores. Pearson correlates were also used to determine the

relationship between each subtest of the DPST and each factor of the HERS. These individual correlations were placed in the cells of a 5x5 matrix. The relationship between total HERS scores and total CTBS scores was determined. Pearson correlations were also placed in the cells of a 5x5 matrix which displayed the relationship between each subtest of the CTBS and each factor of the HERS. The Pearson product-moment correlation was chosen because all of the instruments to be used have at least "equal appearing" intervals. A Pearson product-moment correlation analysis calculates the degree, direction, and significance of relationship between two sets of data.

The Step-wise multiple regression was intended to determine the combination of predictor variables (subtests on the DPST) which best predicted the criterion variable (the HERS). The multiple regression technique finds the best linear prediction equation and evaluates its prediction accuracy, and control for other confounding factors in order to evaluate the contributions of specific variables.

CHAPTER IV
RESULTS

Results of the data analysis of this study are presented in the following manner:

1. Predictive validity analysis of the Dallas Preschool Screening Test with the Hillsborough Early Rating Scale.

2. The multiple regression analysis of Dallas Preschool Screening Test and Hillsborough Early Rating Scale.

3. Criterion comparison of the Hillsborough Early Rating Scale and the Comprehensive Test of Basic Skills.

4. Additional analysis of patterns of extreme scores.

Predictive validity analysis

One hundred and one subjects were compared on a predictor and a criterion variable administered one and one-half years

apart. The DPST, a measure of developmental readiness, was administered during the preschool year and the HERS, a measure of school success, was completed at the end of the kindergarten year. It was hypothesized that performance on the DPST would correlate significantly with ratings on the HERS.

Pearson product-moment correlations were used to compare the two measures. Only five of the 36 correlations were significant at or beyond the. 05 level. The DPST Motor subtest had negative correlations at the. 05 level of significance with the HERS Motor Development factor and HERS total, and correlated negatively at the. 01 level of significance with the HERS factors of Language Development, Cognitive Functioning, and School Achievement (see Table 2). Although five correlations were statistically significant at the. 05 level, or better, the amount of variance held in common was so small that predictive ability was not demonstrated. Performance on the DPST did not predict ratings on the HERS.

Multiple Regression

The multiple regression analysis was attempted but not performed. Once the first variable was entered into the regression formula, no other variables met the requirement of significantly increasing the prediction, so the analysis was terminated.

Criterion comparison

The CTBS was administered to the 101 kindergarten students in the Spring of 1983. The HERS was completed by the students' teachers approximately one month later. Pearson product-moment correlations were used to compare scores on these two measures. Table 3 shows the matrix of correlations between the CTBS and HERS. Twenty-five of 36 correlations were significant at the. 05 level, or better. The CTBS Vocabulary subtest was the only CTBS subtest not to have a significant relationship with any of the HERS factors and total. Sound Recognition and Oral Communication were the CTBS subtests most highly correlated with the HERS total. Despite statistical significance these correlations reflect only a minimal predictive relationship between the CTBS and the HERS.

Table 2. Pearson Product-Moment Correlations of DPST x HERS

HERS	DPST					
	Psych	Aud	Vis	Lang	Motor	Total
Mental Health	-.023	-.028	.163	.075	-.088	.026
Language Development	.033	-.011	.096	.107	-.262**	.002
Cognitive Functioning	-.074	-.135	.085	.028	-.243**	.083
Motor Development	.036	-.039	.113	.139	-.174**	.028
School Achievement	-.109	-.137	.066	.129	-.262**	-.101
Total	-.033	-.081	.127	.081	-.202*	-.022

N = 101

one-tailed tests

* = significance at the .05 level

** = significance at the .01 level

Table 3. Pearson Product-Moment Correlations of CTBS x HERS

HERS	CTBS					
Recog.	Visual Recog.	Sound	Vocab. Comm.	Oral. Total	Reading	Math
Mental Health	.131	.132	-.010	.209*	.157	.166+
Language Development	.194*	.297**	.069	.264**	.235**	.175
Cognitive Functioning	.226*	.301**	.046	.237**	.204*	.118
Motor Development	.296**	.227*	.127	.223*	.229*	.253**
School Achievement	.174*	.292**	.029	.230	.190*	.147
Total	.220*	.255**	.037	.248**	.208	.181*

N = 101

one-tailed tests

* = significant at the .05 level

** = significant at the .01 level

Additional analysis

Due to the inconclusiveness of the statistical analysis, an additional analysis of extreme scores was undertaken. Data considered included low DPST, low HERS, high DPST, and high HERS scores. The only pattern or tendency detected was that low functioning students, as measured by the DPST, scored near the mean one and a half years later on both the HERS and the CTBS.

In addition to low or high total scores, the extreme scores were analyzed for patterns of race, sex, age and "high", "middle" or "low" class ranking. No patterns emerged for race, sex or relative class ranking. It was noted that the two poorest HERS scores were obtained by students in the youngest quartile of the sample. Of the nine lowest students on the HERS, four fell in the youngest quartile. Of the eight highest students on the HERS three were in the youngest quartile. Age does not appear to be a substantial factor in explaining extreme scores.

CHAPTER V
DISCUSSION AND RECOMMENDATIONS

The purpose of this study was to estimate the ability of the Dallas Preschool Screening Test to predict a child's success in kindergarten as measured by the Hillsborough Early Rating Scale, and to compare the HERS to the Comprehensive Test of Basic Skills, a standarized achievement test. It was hypothesized that scores obtained on the DPST would predict HERS ratings and that the CTBS and the HERS were not expected to correlate highly because the HERS purports to measure factors other than achievement that contribute to school success such as mental health, social and physical development, and motivation. The CTBS is primarily an academic achievement measure. It was anticipated that together the HERS and the CTBS would be an accurate measure of kindergarten success.

Pearson product-moment correlations were used to compare tests. All subtest and total scores were used. A multiple regression analysis was attempted but not completed. After the first variable was entered into the regression formula, the addition of the remaining variables did not significantly increase the predictive ability of the DPST. This indicated that the DPST explained little of the HERS variance. Results of these data analyses indicated that the DPST had a limited predictive relationship with the HERS and that the CTBS had only a minimal predictive relationship with the HERS.

Small but significant negative correlations were found between the DPST Motor subtest and the HERS. This finding suggests that children who performed well on the DPST Motor subtest were not rated highly by their kindergarten teachers. An informal item analysis of the DPST Motor subtest indicated that the gross motor items accounted for the variance in scores. Thus, those students who did well on the gross motor items were rated poorly on the HERS.

Interpretations of this result must consider that children who have advanced gross motor skills are likely to be highly active (Hetherington and Parke, 1979). Active children are often perceived by teachers as distractible and thus, receive poor

ratings (Brophy and Good, 1974). Another consideration is that variability in motor development is a function of child rearing practices and stimulations (LaBarba, 1981). Motor superiority has been attributed to greater permissiveness in child rearing practices. Parents who are more permissive and who stress motor stimulation are less likely to stress intellectual stimulation or have an academic achievement orientation. This would help to explain the low HERS ratings of children with high DPST motor scores.

It is notable that the correlations between HERS factors and DPST total are almost. 00. There is a possibility that the significant negative correlations between HERS factors and DPST Motor subtest may have masked any positive correlations among other DPST subtests; thus, yielding DPST total correlations of. 00. A comparison between DPST total minus DPST Motor and HERS factors may have demonstrated higher correlations.

Except for the DPST Motor subtest, the statistical analysis yielded nonsignificant correlations between the DPST and the HERS. As a result, a visual search for patterns of extreme scores was undertaken. Factors of race, sex, age, and relative class ranking were considered. Relative class ranking refers to whether the teacher rated the entire class as being of "high", "middle", or "low" ability or "some of each" on the HERS. The HERS rates each child using his or her classmates as the standard for comparison. None of the factors provided any new information that would help to explain the poor statistical results of this study.

The major disadvantages of an ex-post facto research design are the effects of confounding variables on the gathered data (Kerlinger, 1979). One important confounding variable in the comparison of DPST and HERS was intervention. The predictive ability of the DPST would be greatly affected by any intervention strategies that might remediate deficit areas. This study controlled for any kindergarten intervention that might have been developed on the basis of the DPST by not distributing DPST scores to kindergarten teachers. Other interventions such as those implemented in the preschool year or those based on the results of kindergarten screenings were not controlled for and may have weakened the results.

A second confounding variable was the sample selection process. The study used students having DPST, HERS, and CTBS scores. These requirements severely restricted the number of available subjects making random selection impossible. It must be considered that the sample may not have represented a normal distribution of abilities.

A third confounding variable was biological maturation. Variability in maturation rates between subjects may have confounded the predictive relationship between the DPST and the HERS. A visual analysis of extreme DPST scores revealed that most of the students having very low DPST scores received average scores on both the HERS and the CTBS a year and a half later. Such an improvement in performance might have been the result of a spurt in development that allowed the low DPST students to "catch up" to their peers.

The limitations of an ex-post facto design might have contributed to the poor relationship between the DPST and the HERS. Correlations between the CTBS and HERS were not affected by the confounding variables of intervention or maturation because these two instruments were administered only one month apart. Another explanation for the outcome of this study is problems with the instruments themselves.

The comparison between the DPST and the HERS must reflect the structural differences between the predictor and the criterion. The DPST is a record of a child's actual performance in a given set of items at one point in time, while the HERS is a teacher's interpretation of a child's behaviors in certain areas over a long period of time. These structural differences may have contributed to the lack of significant results. Research about the accuracy of teacher ratings of student behavior has been inconclusive (Brophy and Good, 1974). Studies have both supported and challenged the teacher's ability to objectively rate student performance. A recommendation for improving the HERS involves training teachers to be objective and precise in rating student abilities.

Another difference between the DPST and the HERS is that the DPST, as reported in the manual, measure five factors, while

the HERS measures one factor. The results of this study indicate that the HERS actually measured just one factor such as motivation, interest, or attention. Evidence of this single factor was found in three analyses. First, a closer investigation of the Barnard, Neville and Hilderbrand (1983) study of the HERS showed that although the test items fell into five statistically significant factors, the first factor was so large as to account for most of the variance. The other four factors accounted for only a small percentage of the variance. Second correlations among the HERS factors were high, ranging from. 60 to. 93. This suggests that the subtests measured much of the same attribute. Finally, a visual analysis of item variance within each subject indicated only slight variability in ratings. Teachers were not rating the children's abilities on each discrete item but on the one large attribute.

This interpretation of the HERS data supports the explanations cited for the negative correlations between HERS factors and the DPST Motor subtest. Teachers may have rated those children with advanced gross motor skills based on their activity level or their lack of achievement orientation, not on their abilities in 32 discrete areas.

Correlations among DPST subtests were moderate, ranging from. 17 to. 88, with a median of. 48. CTBS correlations ranged from. 23 to. 90, with a median of. 47. Each of these tests showed more moderate correlations between its subtests indicating that the subtests may have measured discrete attributes. Another hypothesis must consider that the subtests were measuring large amounts of error.

The evidence that the HERS measured one attribute helps to explain the statistical outcome of this study. If the HERS measured a social factor, it would not be expected to correlate highly with a measure of developmental readiness such as the DPST or a measure of academic achievement such as the CTBS. The DPST is administered on a one-to-one basis so that the student's performance is not greatly influenced by motivation or skills of concentration. The CTBS does require more independent work since it is administered in a small group. Motivation and concentration would more likely affect the

student's performance on that test. Hence, the CTBS may have measured some of the HERS attribute, resulting in a small correlation.

An additional consideration of the differences between the tests is that the standards for comparison differ from test to test. The DPST and CTBS use normative samples, but the HERS compares each student to the other students in his or her class. Therefore, different standards were used for each class. It is questionable whether such diverse standards can be compared with any accuracy.

The problems involved in contrasting two tests might have contributed to the outcome of this study but problems within each individual test must also be considered. One weakness of the HERS was the teachers' tendency to rate all items similarly for each child. In this study poor variability between items was particularly true at the positive end of the scale. Teachers did not rate a child on each individual item but on the one large factor previously discussed.

Teachers also tended to choose ratings at the higher end of the scale. On a scale of one to five, with one highest, the overall mean rating was 2.2. In other words, the total distribution of scores was skewed. This restricted range of HERS scores would curtail the relationship between it and both the DPST and the CTBS.

Structural aspects of the DPST must also be discussed as contributing to the nonsignificance of the statistical correlations. The DPST attempts to measure very broad attributes with a small number of items. This is most evident in the Motor subtest which has three items made up of seven tasks. The Language subtest has three items composed of nine tasks. The possibility or error is greatly enhanced by the limited number of items.

Graphs of scores on each subtest of the DPST revealed an additional problem with the instrument. Three scales, Auditory, Visual, and Motor, were not continuous but contained three or

four specific scores on which most points fell. Such distributions would confound the relationship between the DPST and the HERS.

Problems within each individual test as well as the difficulties involved in comparing tests and confounding variables have been discussed as possible explanations for the results of this study which indicated that the DPST had virtually no meaningful predictive relationship with the HERS. No significant relationship is generally interpreted as an indication that the two instruments do not measure the same attributes. This was the case between the DPST and the HERS. There is substantial evidence that the HERS does not measure academic achievement nor sensory/motor development. The HERS probably measures something quite different: a child's motivation, interest in school, skills of concentration, or simply "looking good" to a teacher. It is therefore, not surprising that the DPST and the HERS did not correlate. One of the purposes of this study was to determine if the DPST could predict success in kindergarten. The HERS may be an indicator of school success but the vast differences between the attributes being measured, as well as other problems between and within the two tests have significantly weakened it's power as a criterion variable.

The other purpose of the study was to compare the CTBS to the HERS. These two tests had a minimal predictive relationship. The CTBS is more similar to the DPST in it's content and structure. The greater correlation between the CTBS and HERS than between the DPST and the HERS was most likely due to the relative concurrency of the administration of the former. The fact that the CTBS and HERS were administered only one month apart yet had such small relationships suggest that, individually, they are not both good measures of school success. The CTBS may measure achievement and the HERS may measure motivation but independently they may not amply measure school success. Achievement and motivation might be important components of school success but not sufficient one without the other. A variety of factors have been indicated as origins of problems in this study. It is an impossible task to find

the precise cause of the poor results. Weaknesses in each of the tests must be considered as primary sources of error. As a predictor, the DPST remains questionable. More research is necessary before it can be used with confidence as a preschool screen.

Despite the nonsignificance of the correlations between the DPST and HERS, preschool screening is still considered to be important. The process of identifying potential deficit areas and providing remediation before a child begins his or her school career can be an influential component of a preschool program. Many studies have had significant and meaningful results using preschool screens to predict later school success (Conrad & Tobeissen, 1967), Frankenburg et al, 1971, and Satz, 1982). The validity of the DPST will be strengthened with the improvements of more complete directions and local norms.

The HERS was chosen as the criterion variable because it was intended to measure five different attributes both academic and social/emotional. One interpretation of the results of this study is that the HERS, in fact, measures one primarily social attribute. It is hypothesized that this attribute is an important component of school success and that it will remain stable over time. If this were true, then a child who received a high rating in kindergarten should also score highly on the HERS in first grade. It is recommended that future research involve a test-retest reliability study of the HERS. Such a study would provide evidence of the stability of the instrument over time and would support the hypothesis that it measures an important component of school success.

This study has contributed to a greater understanding of the DPST, and particularly, the HERS. It is proposed that the CTBS and the HERS measure distinct yet equally important aspects of school success. A final recommendation for future research in the area of preschool screening is to combine assessment devices to include both academic and social measures. This would be advantageous for both the predictor and the criterion variables. It is expected that the two measured together will provide a more accurate picture of a child's success in school.

In conclusion, this study has made specific contributions to educational institutions that are involved in the screening of preschoolers. First, the results of this study do not support the use of the Dallas Preschool Screening Test as a predictor of later school success as measured by the HERS. The continued use of the DPST in this capacity is questionable. Second, this study has provided evidence that the Hillsborough Early Rating Scale, in its present form, does not measure five factors contributing to school success but one social factor. Finally, the results of this study indicate that combining measures of achievement with measures of social functioning will produce more accurate estimates of school success.

REFERENCES

Adelman, H.S., and Feshbach, S. (1971). Predicting reading failure beyond the readiness model. *Exceptional Children, 37,* 349-354.

Attwell, A.A., Orpet, R.E., and Mayers, C.E. (1967). Kindergarten behavior ratings as a predictor of academic achievement. *Journal of School Psychology, 6*(1), 43-46.

Barnard, J., Neville, D. and Hilderbrand, J. (1983). *The development of a teacher's rating scale to determine school performance in young children.* Unpublished report, University of South Florida, Tampa.

Brigance, A. (1981). *Brigance Inventory of Early Development.* North Billerica, Ma.: Curriculum Associates.

Bronfenbrenner, U. (1974). *A Report on Longitudinal Evaluations of Preschool Programs: Volume II: Is Early Intervention Effective?* Office of Child Development (DHEW), Washington, D.C. (ERIC Document Reproduction Service No. ED 093501).

Brophy, J. and Good, T. (1974). *Teacher-Student Relationships Causes and Consequences.* New York: Holt, Rinehart and Winston, Inc.

Bryan, T.S., and McGrady, H.J. (1972). Use of a teacher rating scale. *Journal of Learning Disabilities, 5,* 199-206.

Casey, J.M. (1976, January). *A pre-kindergarten screening.* Unpublished doctoral dissertation. (ERIC Document Reproduction Service No. ED 145966).

Conrad, W.G., and Tobeissen, J. (1967). The development of kindergarten behavior rating scales for the prediction of learning and behavior disorders. *Psychology in the Schools, 4*(4), 359-363.

Daniels, M., and Schouten, J. (1970). *The Screening of Students.* London: G. Harrup and Co. Ltd.

Davis, F.B. (1974). *Standards for Educational and Psychological Tests*, Washington, D.C., American Psychological Association, Inc.

Devoid, R.R. (1976). The predictive validity of a multi-model preschool program. *Dissertation Abstracts International.* 37(8), 4970-A.

Frankenburg, W.K., Camp, B.W., and Van Natta, P.A. (1971). Validity of the Denver Developmental Screening Test. *Child Development, 42,* 475-485.

Glazzard, M. (1977). The effectiveness of three kindergarten predictors for first grade achievement. *Journal of Learning Disabilities, 10,* 95-99.

Glazzard, P., Tollefson, N., Selders, J., and Barke, C.R. (1982). The predictive validity of the kindergarten teacher rating scale. *Educational and Psychological Measurement, 42,* 687-93.

Haring, N.G., and Ridgeway, R.W. (1967). Early identification of children with learning disabilities. *Exceptional Children, 33*(6), 387-395.

Hetherington, E.M. and Parke, R.D. (1979). *Child Psychology A Contempory Viewpoint.* New York: McGraw-Hill Book Company.

Hinton, G.G., and Knights, R.M. (1971). Children with learning problems: academic history, academic prediction, and adjustment three years after assessment. *Exceptional Children, 37,* 513-519.

Hockman, T. (1972). *Focus on preschool developmental problems.* Colorado Springs: Colorado Public Schools (ERIC Document Reproduction Service No. ED 078945).

Keough, B.K., and Becker, L.D. (1973). Early detection of learning problems: questions, cautions, and guidelines. *Exceptional Children, 40,* 5-11.

Keough, B.K., and Smith, C.E. (1967). Visuo-motor ability for school prediction: a seven-year study. *Perceptual and Motor Skills, 25,* 101-110.

Keough, B.K. and Smith, C.E. (1970). Early identification of educationally high potential and high risk children. *Journal of School Psychology, 8*(4), 285-290.

Kerlinger, F.N. (1979) *Behavioral Research: A conceptual approach.* New York: Holt, Rinehart and Winston, Inc.

Koppitz, E.M., Sullivan, J., Blyth, D.D., and Shelton, J. (1959). Prediction of first grade school achievement with the Bender-Gestalt test and human figure drawings. *Journal of Clinical Psychology, 15,* 164-168.

Kresh, E., and Green, B.F. (1969). *Preschool academic skills test.* Paper presented at the American Educational Research Association meeting, Los Angeles. (ERIC Document Reproduction Service No. ED 028045).

LaBarba, R. (1981). *Foundations of Developmental Psychology.* New York: Academic Press.

Lambert, D.M. (1980). The reliability and concurrent validity of the Dallas Preschool Screening Test. *Dissertation Abstracts International, 41,* 930-A.

Lichtenstein, R. and Ireton, H. (1984). *Preschool screening. Identifying young children with developmental and educational problems.* Orlando: Grune and Stratton, Inc.

Luce, S.R., and Hoge, R.D. (1978) Relations among teacher rankings, pupil-teacher interactions, and academic achievement: a test of the teacher expectancy hypothesis. *An Educational Research Journal, 15*(4), 489-500.

Mardell, C., and Goldenberg, D. (1983). *Developmental Indicators for the Assessment of Learning - Revised Manual.* Edison N.J.: Childcraft Education Corporation.

Matheny, A.P. (1971). Comparibility of WISC and PPVT scores among young children. *Exceptional Children, 38,* 147-150.

Meisels, S.J. (1976). *Developmental screening in early childhood.* Washington, D.C.: The National Association for the Education of Young Children.

Metham, J. (1978). *Teacher's ratings of preschool children's behaviors.* (Discussion paper number 2). North Ryde, Australia: Macguire University, School of Education. (ERIC Document Reproduction Service No. ED 161542).

McKnab, P.A., and Fine, M.J. (1972). The Vane kindergarten test as a predictor of first-grade achievement. *Journal of Learning Disabilities, 5*(8), 52-59.

Nunnally, J.C. (1978). *Psychometric Theory, 2nd ed., New York: McGraw-Hill Book Co.*

Percival, R.R. (1972). *Dallas Preschool Screening Test Manual,* Richardson, Texas: Dallas Educational Services.

Reynolds, C.R. (1979). Should we screen preschoolers? *Contemporary Educational Psychology, 4,* 175-181.

Rosenberg, L.A., and Stroud, M. (1966). Limitations of brief intelligence testing with young children. *Psychological Reports, 19,* 115-118.

Rubin, R.A., and Balow, B. (1968). *A comparison of pre-kindergarten and pre-first grade boys and girls in measures of school readiness and language development.* (Interim Report) Minneapolis: University of Minnesota. (ERIC Document Reproduction Service No. ED 023474).

Satz, P., and Fletcher, J. (1982). *Manual for the Florida Kindergarten Screening Battery,* Odessa, Florida: Psychological Assessment, Inc.

Segerstrom, R.R. (1976). A study in the prediction of school readiness and school failures. *Dissertation Abstracts International, 37*(11), 7050-A.

Silver, A.A. (1977). *Basic Handbook of Child Psychiatry,* New York: Basic Books, Inc.

Southworth, L.E., Burr, R.L., and Cox, A.E. (1980). *Screening and Evaluating the Young Child,* Springfield, Illinois: Charles C. Thomas Publishers.

Virgin, A., and Schecter, H. (1974). *Accuracy of kindergarten teachers' predictions of their pupils' subsequent performance.* (Research Report). Willowdale, Ontario: Board of Education. (ERIC Document Reproduction Service No. ED 130766).

Wendt, R.N. (1978). Kindergarten entrance assessment: is it worth the effort? *Psychology in the Schools, 15*(1), 56-62.

Wendt, R.N. (1979). Pre-kindergarten screening point-counterpoint. *Viewpoints in Teaching and Learning, 55*(3), 18-24.

APPENDIXES

APPENDIX 1

HILLSBOROUGH EARLY RATING SCALE

Student's Name_____

TEACHER'S SURVEY DIRECTIONS

The name of the student you are to evaluate appears at the top of each survey. Each item in the questionnaire contains key words or phrases (underlined) and associated definitions that you are to use to rate each student. Please compare the student with other students in his/her class. The information collected will be held in strict confidence.

The rating scale is a five-point scale. A rating of "1" indicates the student is among the top two or three students in his/her class; a rating of "2" means the student is in the upper middle group in his/her class; a rating of "3" means the student is in the middle group in his/her class; a rating of "4" means the student is in the lower middle group of his/her class; and a rating of "5" means the student is among the lower two or three students in his/her class.

BEST COPY AVAILABLE

Teacher Name: _____ (optional) Date: _____

Student Name: _____ Sex: Male ___ Female ___

Race: Black __ White __ Hispanic __ Other __ Date of Birth: _____ Month Day Year

Type of class this child is in (check one):

high ability ___ middle ability ___ low ability ___ some of each ____

Total number of students in the class _____

Rating

1 = "top two or three students"
2 = "upper middle group"
3 = "middle group"
4 = "lower middle group"
5 = "lower two or three students"

	Characteristics	Rating (circle one)				
1.	BRIGHTNESS: comprehends and learns rapidly and easily	1	2	3	4	5
2.	THINKING SKILLS: the ability to reason, solve problems, generalize, see relationships, and make logical associations	1	2	3	4	5
3.	RETENTION: retains what is learned without much drill	1	2	3	4	5
4.	SELF-ORGANIZATION: efficient in coordinating the materials/ time required for independent task completion	1	2	3	4	5
5.	DESIRE TO EXCEL: motivated to learn and consistently puts forth maximum effort	1	2	3	4	5
6.	LISTENING COMPREHENSION: understands teacher verbalization, especially directions	1	2	3	4	5
7.	VOCABULARY: rich vocabulary marked by originality of thought/expression	1	2	3	4	5
8.	SENTENCE STRUCTURE AND LENGTH: mature quality of each	1	2	3	4	5
9.	MATURITY OF PRONUNCIATION: grade appropriate articulation	1	2	3	4	5
10.	ORAL COMMUNICATIONS SKILLS: the quality of relating thoughts and ideas	1	2	3	4	5
11.	FINE MOTOR COORDINATION: drawing mature and detailed, consistent hand preference with acceptable pencil grip	1	2	3	4	5
12.	GROSS MOTOR: good overall coordination including hopping, skipping, and jumping	1	2	3	4	5

BEST COPY AVAILABLE

BEST COPY AVAILABLE

	Characteristics	Rating (circle one)				
13.	FINE MOTOR COORDINATION: copying shapes, letters, prints name, and cuts with scissors	1	2	3	4	5
14.	SELF-RELIANCE AND INDEPENDENCE: separates readily from parents, manages self-help tasks, requires little support in task completion	1	2	3	4	5
15.	EMOTIONAL MATURITY AND SELF-CONTROL: behaves in a mature manner and conforms to classroom rules	1	2	3	4	5
16.	GROUP PARTICIPATION: frequent volunteer, complies without reluctance	1	2	3	4	5
17.	LEADERSHIP QUALITIES: obvious leader among classmates	1	2	3	4	5
18.	SENSE OF HUMOR: comprehends teacher jokes/puns, makes clever/humorous remarks	1	2	3	4	5
19.	POSITIVE ATTITUDE TOWARD SCHOOL: appropriate work habits, responsive, interested	1	2	3	4	5
20.	POSITIVE ATTITUDE TOWARD PEERS: socializes appropriately with classmates needing little teacher intervention	1	2	3	4	5
21.	POSITIVE ATTITUDE TOWARD SELF: cheerful, relaxed, confident	1	2	3	4	5
22.	ADAPTABILITY: demonstrates appropriate reaction to change in daily routine	1	2	3	4	5
23.	ATTENTION DURING STORYTIME: enjoys storytime, pays close attention, good attention span	1	2	3	4	5
24.	EXPRESSED DESIRE TO READ/WRITE OR HAS STARTED TO DO SO: eager to learn	1	2	3	4	5
25.	CURIOSITY: interested in diverse topics	1	2	3	4	5
26.	ENJOYS BOOKS/WORDS/IDEAS: frequently chooses to look at books, asks for definitions/explanations, enjoys expanding ideas	1	2	3	4	5
27.	ACCEPTANCE BY CLASSMATES: establishes friendships, shares materials	1	2	3	4	5
28.	PHONICS: knows letter sounds, blendings	1	2	3	4	5
29.	KNOWLEDGE OF NUMBERS: counting, recognizing number symbols	1	2	3	4	5
30.	PROPERTIES OF ARITHMETIC: knows concepts (larger, smaller, etc.)	1	2	3	4	5
31.	KNOWLEDGE OF ALPHABET/WORDS: comes to school knowing most or all of the alphabet and some or many words	1	2	3	4	5
32.	OVERALL, your evaluation of this student is	1	2	3	4	5

SB88053

BEST COPY AVAILABLE

APPENDIX 2

TABLE 4

MEANS AND STANDARD DEVIATIONS OF DPST, HERS AND CTBS

DPST	Mean	Standard deviation
Psychology	41.38	9.19
Auditory	31.57	11.14
Visual	46.34	7.73
Language	38.76	7.24
Motor	18.05	5.19
Total	176.10	31.30
HERS		
Mental Health	18.46	7.41
Language Development	9.43	3.87
Cognitive Functioning	13.38	6.03
Motor Development	6.81	2.76
School Achievement	6.00	2.81
Total	72.48	27.10
CTBS		
Visual Recognition	503.21	83.41
Sound Recognition	448.20	79.74
Vocabulary	367.52	73.29
Oral Communication	426.48	127.39
Reading Total	397.41	82.69
Math	441.06	88.01

SUBANAVA